The Future of Forth

Forth on a Chip

Forth on a Mac: MacFORTH & Neon

nroff Hyphenation on a Micro

MSDOS Quirks

Hobbles on Handles

y Pac™ For The Holidays.
5. $995.*

16-bit processors is supported, along with the user-assignable string and numeric variables, up to 100-step trace-back, and debug log to disk or printer. Breakpoints can be set in the source file window, whose display can be synchronized with the disassembled object during tracing; symbols can be added incrementally during the debug session; and the disassembly written to disk.

…Pfinish,™ the program that maximizes your program's efficiency…

Pfinish helps you turn your beta-test product into a software work of art. It analyzes your program or the entire operating environment during execution, and produces reports which tell you how much time was spent in each routine or interrupt, who called it, how many times, and much more. Wasteful, inefficient, or non-optimal areas of code become immediately apparent, whether they're in your code, the compiler library, or in the operating system routines themselves.

…Ptel™ gets the lead out of binary file transfer…

You get error-free file transfer and access to mainframes, minis, and micros. ASCII. XModem. Modem7. Telink and Kermit. MS-DOS® 2.x or 3.x. With Ptel. You can transfer 8-bit binary files over a 7-bit data path with Kermit. Masked file-name transfers with XModem. Or, transfer whole subdirectories with a single command using Telink. Ptel keeps track of the original file size and creation date. Ptel also offers a script language, backward scrolling, and the ability to handle DOS commands from inside the program.

…And, Phoenix's new on-line software update service. Free to all registered Pfantasy Pac1 owners.

Pfantasy Pac1 software works in MS-DOS/PC DOS environments, is completely language-independent, and is available for the IBM® PC, XT,™ AT,™ and compatibles, the Wang® PC, the TI Business-Pro,™ and the Tandy® 2000.

So, spoil yourself a little. Get Pfantasy Pac1, and make your favorite programmer's dream come true.

For more information contact: Phoenix Computer Products Corporation, 320 Norwood Park South, Norwood, MA 02062. Or call: (800) 344-7200. In Massachusetts (617) 762-5030.

Programmers' Pfantasies™
by

Phoenix

SEE US AT FALL COMDEX

CONTENTS

In This Issue

Once again we present our annual Forth issue. Perhaps the most significant development in Forth over the past year has been the direct implementation of Forth engines in silicon logic. One such silicon Forth engine is the Novix NC4000 chip. In our feature article Leo Brodie describes some of the revolutionary features of this chip and the benefits that it offers to programmers.

This month we also provide a solution to the problems of developing tight, rapidly executing code for PROM-based applications written in Forth. H. Robinson, P.D. Morse and S.A. Bowhill of the Central Illinois Chapter of the Forth Interest Group have designed a post-compiler that selects from compiled, threaded Forth code the segments necessary for the execution of the PROM-based application and redirects to the RAM of the target machine any code that must be resident there.

Our review section this month also focuses on Forth, in particular, on Forth and the Macintosh. First of all, Bruce Horn examines Neon, an object-oriented programming environment that is "a cross between Smalltalk and Forth." Then, Alan Clute takes a look at MacFORTH, a development package designed to ease the writing of programs in the Mac environment.

Our columnists have added their usual monthly goodies. Dave Cortesi serves up some further tidbits from his study of MSDOS, showing how to handle ambiguous pathnames with DOS functions 4Eh and 4Fh, how to define MSDOS diskette formats and how to control Control-Z.

Allen Holub had hoped to code a version of the original T_EX hyphenation algorithm in C for last month's issue, in which we reviewed two implementations of T_EX for the IBM PC. Limitations on his time prevented him from realizing his intentions before deadline, but he did manage to finish off the project for this month's C Chest. So, if you want to find out how to hyphenate supercalifragilisticexpialidocious . . .

Bob Blum seems to be gleefully drowning in the deluge of new CP/M products that developers are sending to him. This month he finishes his discussion of the AMPRO Little Board, reviews SLRMAC, the 8080 assembler from SLR Systems, and offers some enticing glimpses of Date Stamper from Plu*Perfect Systems, DSD80, the debugger from Soft Advances and the latest single board computer from Micro Mint.

Finally, we continue this month in Mac Toolbox the program listings for the MacSCSI hard-disk interface presented by John Bass last month. Look for more on the MacSCSI next month.

Dr. Dobb's Journal

EDITORIAL

Charles Moore first created Forth to control astronomical instruments; lately his attention has been focused on the design and implementation of a microprocessor that uses Forth as its machine language. The Novix chip represents only one recent effort to put Forth on silicon. Rockwell fit a good chunk of Forth into its R65F11 chip, along with the 6502 instruction set. The British Metaforth MF16LP single-board computer, which is implemented with custom bipolar circuits, uses Forth as its machine language. Bipolar technology is also incorporated in the H4TH/X Forth engines from Hartronix, which provide a real-time system with up to 4,000 primitives in firmware.

None of these products (with the exception, perhaps, of the H4TH/20) is of much practical significance to the Forth programmer developing commercial software on and for personal computers. They are more likely to interest programmers working in areas similar to those worked on by the first Forth programmer: device control, process control, robotics. Forth on a chip can only help speed developments in these fields.

There are grounds for wondering whether the programmer who uses Forth to develop personal computer applications is in a dwindling minority. The virtues of Forth in such applications are similar to those of C. Are Forth's additional virtues the kind that will secure it a permanent place in personal computer software development? Perhaps they are, but doubts in that area may have been one of the reasons why the language Neon was developed as an extension of Forth. Even so, extending Forth to the point where it becomes a different language is not a new idea.

Forth seems in excellent health, but hardware Forths and extended Forths lead one to wonder in just what area Forth will see its greatest use in the next year, and even whether the language will soon cease to be regarded as a high-level language. One wonders: Where is Forth bound?

Where *DDJ* is bound is, we hope, evident from its ten-year trajectory: on a path of providing more software tools for advanced programmers. We can be more specific. Here is our 1986 editorial calendar.

January: 68000 Programming. Article deadline: past.
February: Pascal, Modula-2, Ada. Article deadline: now.
March: The Future of Programming: algorithms for parallel processing. Article deadline: 12/1/85.
April: Artificial Intelligence. Article deadline: 1/1/86.
May: At the Human Interface: designing usable software. Article deadline: 2/1/86.
June: Communications. Article deadline: 3/1/86.
July: Forth. Article deadline: 4/1/86.
August: C. Article deadline: 5/1/86.
September: Algorithms. Article deadline: 6/1/86.
October: 80286/80386 Programming. Article deadline: 7/1/86.
November: Graphics. Article deadline: 8/1/86.
December: Operating Systems. Article deadline: 9/1/86.

Plus, of course, our regular columns on C, Unix, 16-bit software, CP/M and Macintosh software development tools.

Michael Swaine

Michael Swaine

Hardware

Dear *DDJ*,

I vote for more and continuing hardware coverage, construction and otherwise. Every magazine except *BYTE* seems to be doing less and less with it. Yet, over time, Steve Ciarcia's hardware features are consistently among the most popular *BYTE* publishes.

The print spooler [in the July, 1985 issue of *DDJ*] can be simplified considerably if 6264 or similar 8K × 8 static RAM is used instead of 4164s. Seven 6264s would give 56K of RAM, leaving the top 8K of the 64K address space for ROM and the 6821 PIA. A 74LS138 ³⁄₈ decoder could break the address space into 8K chunks and supply the enable signals for individual devices. Additional decoding would be required only in the top 8K, containing EPROM and the PIA. That could be done with a single NOR gate. Two 74LS541 buffer packages could drive A0–A12, R/W, and the enable (phase 2) clock. This buffering is desirable due to the loading imposed by the additional 3-state outputs that use of 7 6264s would involve. Software could be much simpler because of elimination of the refresh routines. System clocks would also be simplified. Total cost of construction wouldn't be much more than the spooler presented.

The S-100 real-time clock article was interesting. However, there were a few errors and omissions that deserve comment.

The use of the MM58167A's RDY line to induce WAIT states depends on rather critical timing relationships among various events. These timing considerations can be greatly reduced if an interface adapter such as the 8255 is used to control the clock chip. The only critical timing relationships that exist then are those involving the 8255's interface with the bus. Since it is much faster than the 58167 and its timing requirements more closely matched to system timing, fewer problems occur.

The POWER DOWN input needs to go low 1 microsecond before the power off only if the standby interrupt is connected to something. If this particular feature of the MM58167A clock chip is not being used at any time, the POWER DOWN input can be tied directly to +5 volt power. The 2N2222 and 2N2907 transistors and associated components can then be eliminated. In many cases the power circuit for pin 24 of the clock chip need be no more complex than two diodes, a resistor, a capacitor and a battery; however, if the clock is on a board that tends to be noisy, a more complex zener diode and transistor switch can be used there. An example of this can be seen in Steve Ciarcia's May 1982 *BYTE* article.

While the fastest fixed-interval interrupt frequency specified for the 58167 by National Semiconductor is 10 Hz, a 100 Hz interrupt can be obtained by selecting the counter-latch match (sometime called "alarm clock") interrupt address, then writing "don't care" states into the seconds/tens-of-seconds latches. A good deal of control can be obtained by writing valid comparison values into some of the other latches, and "don't care" into others. A "4" in the day-of-the-week latch and "don't care" everywhere else would get you 100 Hz interrupts all day every Wednesday. A "12" in the hours/tens-of-hours latches and "don't care" elsewhere gets you 100 Hz from 12:00:00 through 12:59:59 every day. Add the "4" in day-of-the-week and you get the 12:00:00 through 12:59:59 100 Hz interrupts every Wednesday. Other combinations of valid and "don't care" states get other patterns.

For most people's purposes, this capability is, at best, somewhat interesting, but otherwise useless. For automatic data logging, however, it could be quite useful.

The GO command increments the counter if the seconds count is greater than 39, not 40, as stated on page 64.

It is frequently easier and more foolproof to use the seconds/tens-of-seconds counter reset address to start the clock at a known time. I think maybe the GO command was either dreamed up because there was an address that wasn't otherwise used, or else it happened by accident. A goof in a mask, perhaps? A bug that became a feature?

On page 62 it is stated: "Each January . . . the clock IC does not automatically change to a new year." While true, it is also misleading: the 58167 doesn't have years/tens-of-years counters (or latches). Any changing of the year gets done somewhere other than the 58167 clock circuit.

Someone who wanted to make a clock board that does have years counters could use an OKI MSM5832, National Semiconductor MM58274, or Motorola MCM146818 or MCM146819 or any other clock chip that tracks years.

Sincerely,
Frank Kuechman
8113 NE 25th Avenue
Vancouver, WA 98665

DDJ

Circle **no. 112** on reader service card.

DR. DOBB'S CLINIC

by D.E. Cortesi

Here's a batch of small items on MSDOS and PCDOS. The first is based on a note from Nick Hammond of Torrens, Australia. There, in a magazine called *Your Computer* and in a column not unlike this one, appeared the plaint of a puzzled user. Sometimes, he said, his everyday programs would misbehave strangely. DEBUG would bomb or fail to write a good file; FORMAT would loop or report track zero bad when it wasn't; and so forth. The only common factors were that the failures occurred when there were resident programs in the system and that they could be fixed by booting without the resident programs. Sometimes they could be fixed merely by changing the order in which resident programs were loaded.

Dismal Memory Access

Hammond wrote to explain this puzzle. When IBM designed the PC, he tells us, they fitted a Direct Memory Access (DMA) chip from an 8-bit family. It can address only 64K. To accommodate the 1Mb address space of the 8088, they tacked on an external "page register" to establish which of 16, 64K "pages" the DMA chip would use. Because an 8088 "paragraph" is 16 bytes and a "page" is usually 256 bytes, we might do better to call these 64K DMA blocks, "chapters."

What we must never call them, Hammond says, is "segments." Unlike the 65K segments the 8088 can address, which may begin at any paragraph boundary, there are only 16 DMA chapters, and they fall exactly on the 64K boundaries of storage.

As a result, the DMA chip is physically incapable of transferring a block of data that crosses a chapter boundary in a single operation. "You have to transfer as far as the boundary, change the page register, then transfer the rest," says Hammond. "This would be fine if the firmware in the BIOS ROM handled the housekeeping, but it doesn't."

For example, if you use BIOS interrupt 13h to access a floppy disk and your data area crosses a chapter boundary—if any byte in your buffer except the first has an address that is divisible by 65,536—the BIOS will return an error. Programs that don't specifically check for this condition will fail intermittently, depending on their load address, and that depends on what resident programs were loaded ahead of time.

It's our impression that I/O done through MSDOS does not suffer this constraint—just one more reason to avoid ROM calls.

Control That Zee!

MSDOS function 40h writes up to CX bytes to the file named by the "handle" in BX and returns in AX the number of bytes written. Ordinarily, all the bytes will be written and AX will equal CX afterward.

Usually the reason DOS will report writing fewer than CX bytes is that the output disk is full and there is no place to put them. That's a disastrous error from which a program has few courses of recovery.

We recently discovered another cause—one that isn't documented so far as we are aware. If the output data include a logical end of file byte (Control-Z), and if BX names the standard output handle (0001), and if standard output has *not* been redirected to a file, DOS will stop writing at the Control-Z and return the count of bytes that preceded it.

This behavior makes sense in a way; it makes it easy to copy a disk file to the screen without displaying the superfluous garbage that may follow logical end of file. On that account, we'd expect DOS to behave the same whenever the target file was implemented by a character device driver, whether it was named by a standard handle number or not. We didn't test to see if it was that consistent, or whether it only applied the rule to the handles for the standard output, print, and auxiliary files, or whether it was just a glitch on handle 0001, or whether it varied with the device driver.

The more consistent DOS is, the worse the problems are. How is a program supposed to distinguish a full disk from an incidental Control-Z? The only way we can think of is to use the ioctl function to get the attributes of the file's device driver, and see if it's a block device or a character one. Wonderful.

One program that doesn't check is the MS Pascal 3.0 runtime library. If a Pascal program copies from standard input to standard output and if input is an edited keyboard line including Control-Z and output is the screen, it will abort with a disk-full error.

Medium Smart

Further gleanings from our study of MSDOS can be found in the Table (page 12), a compilation of the numbers that define MSDOS diskette formats. We pieced it together from IBM and Microsoft manuals.

The second line of the table is headed "format flag-byte." That's what we call the first byte of the FAT, the same byte that Microsoft and the illiterati of Boca Raton persist in calling the "media descriptor," thereby driving us straight up the wall. The word "media" is *plural*, dammit, and is in no way a synonym for "disk." The recording *medium* used with dis-

kettes is an iron or chromium oxide in a plastic binder, but the named byte doesn't describe it.

Whatever you call it, that byte is a poor guide to disk layout. The same flags can be found on 5- and 8-inch disks of very different formats, and the flag F9h appears on both 80-track diskettes and fixed disks. There is a better indicator: DOS functions 1Ch and 36h both return in register DX the total number of allocation units on a disk. Read across the third-from-last row of Table 1; you'll see that each format has a unique number of allocation units.

Hear No Evil . . .

. . . is the policy of Microsoft regarding MSDOS. We found a bug in DOS and wrote it up in these words: "DOS function 43h, AL=00 (return file attributes of path *DS:DX) returns impossible values and no error in the carry flag when given the path 'C:\.' Under DOS 3.0, despite absence of a carry error signal, function 59h returns partly-correct, partly-contradictory information, suggesting an error was recognized." We attached debug output to prove these points and mailed it all to Seattle.

It was a waste of time. Microsoft is not interested in trouble reports from civilians; they only talk to companies. Or so said Doug Boa of Microsoft in an earnest reply. "Microsoft does not support either of our operating systems (MSDOS and Xenix) directly," he wrote. "We do not sell them to the public, but only to the equipment manufacturers . . . [who] are responsible for all support functions." Therefore, Boa said, "even though you are reporting a problem with a function call general to all MSDOS implementations, you still need to direct the report to the equipment manufacturer."

We like to think we can read a tone of honest regret in those lines. If we wrote an OS, it would just *kill* us to have to turn back trouble reports because we'd know that, first, the people who report them are more likely to give up than to submit them a second time to their OEMs, and second, the OEM is all too likely to ignore the reports, lose them, reject them out of

ignorance (or habit), or just sit on them for a few months before passing them on.

So let's give a moment's sympathetic thought to the Microsoft developers, sitting around wistfully hoping that somebody will write them a letter they'll be allowed to read. But if you find a bug in MSDOS or Xenix, save your time and postage. We might add that the sysops of the Microsoft forum on CompuServe will also reject any technical queries or trouble reports on MSDOS or Xenix.

Narrow Paths

MSDOS permits a program to control files entirely in terms of compact strings of characters called pathnames. Having to convert the name of a file from a compact, punctuated string like C:DROSS.DAT to the fixed fields of a file control block and back again was a major irritation of file work in CP/M and early MSDOS. With pathnames, programs talk to DOS in the same terms with which users talk to programs. That should make things easier.

And it does, so long as you work only with single files with explicit names. Get what should be a pathname from the user and terminate the input string with a null byte; point to the string and tell DOS, "Open it." DOS takes care of case conversion and of parsing the string into its many parts: drive letter, directory names, filename, extension. If the path describes anything that DOS can accept as a file, your program gets back a numeric "handle" and is off to the races. If it doesn't, DOS returns either error 2, "file not found," or error 3, "path not found." The first means that the disk letter (if any) and directory names (if any) were valid but the filename and extension couldn't be found in the (last-named or default) directory. Error 3 means that something went wrong before the filename and extension could be checked. We should be delighted that DOS can handle all this complexity for us, because it would be a real punctuation (a real #$?!%$*&!) to have to program it.

It's as easy to erase, create, or rename any single file from its path-

name, or set its attributes or time-stamp, as it is to open one.

Winding Paths

Complications arise when you want to handle ambiguous pathnames (ones in which wild cards are used to represent multiple files), especially if you want your program to be as flexible as DIR or COPY.

Let's create an example. Suppose that on drive B you have a diskette with a small hierarchy of files:

```
top.dat
mydir
    inner1.dat
```

inner2.dat

And suppose that at some time you moved to drive B and made MYDIR the current directory:

```
A>b:
B>cd mydir
B>a:
```

That, by the way, is a point that's poorly explained by the DOS manuals we've seen. They don't make it clear that *each* drive has its own current directory. A reference to a file by drive, name and extension, as in the example

A>type b:inner1.dat

means "look up the file named IN-NER1.DAT in the current directory of drive B." It doesn't mean the *root* directory of drive B; that would be

A>type b:\inner1.dat

Nor does it necessarily mean in the directory last named to CD; that might have been a new current directory for another drive.

Now consider all the ways that you can refer to the files on drive B using DIR. You can ask about all files by not supplying any operand,

A>b:
B>dir

which will list all files in the current directory. You can do the same, when another drive's the default, by naming the drive.

A>dir b:

You can ask for all the files in the *root* directory of the disk by specifying "root" with a backslash:

A>dir b:\

By convention of DIR and COPY such a reference to a directory is implicitly a reference to all the files in that directory. It saves your having to type *.*, as in

A>dir b:*.*

or the equivalent double-dot back-reference,

A>dir b:..*.*

You may refer to files in any directory by naming the directories along the path to the files,

A>dir b:..\mydir\..\top.dat

including redundant ones, as in that example.

Primrose Paths
There are a lot of useful programs with specifications that start off like

this: "for every file that matches the first operand, the program will . . ." For consistency, such programs ought to interpret their command operands just as generally as DIR does. Anything less will just confuse the users: they'll come back complaining "DIR finds this file, why doesn't your program?"

The key to handling ambiguous pathnames lies in the DOS functions 4Eh and 4Fh, search for first and next match to path, respectively. Point to a pathname string that may or may not be ambiguous and call DOS; if there's (another) file that matches the pathname, DOS drops information about it into the current data transfer area. One returned item is the filename and extension, formatted as a string.

Unfortunately, functions 4E/4F aren't as general as DIR is. They do not assume *.* after a directory or drive letter, for instance.

Even when they are successful, they leave you with a riddle: how can you *open* the file you've discovered? You have the pathname you used for the search, a user-entered string like b:\mydir*.dat, for example. You have the filename and extension returned by DOS, INNER2.DAT, for instance. Now you'd like to do something to the file you've found—open it or erase it or something—but you don't have enough information. You can't just point to the filename string and request an open. The file may reside in a disk or directory other than the current default, and DOS won't find it. The necessary qualifiers are in the search string, but you can't get at them without parsing the string. Parsing pathnames was the job that we so gladly turned over to DOS.

As you already suspect, we wouldn't be going on about this if we didn't have a solution. We found that we could use DOS function 4300h (return attributes of file using pathname) to tell the type of path a user has provided. With fair reliability, the user's path specification can be processed into two versions: a search path for use with functions 4E/4F, and a lead-in path to be prefixed to the filename string that those functions return.

The logic of it is explained at tedious length in the comments of Listing One (page 16), the source of a C function that does the job. This function, **fixpath()**, would be called once to initialize a path provided by the user. If it returns a zero, there are no files to be found down that path. Otherwise it has made two new paths, a search path and a prefix path. If the input path was, say,

b:..\

("all the files in the parent of the default directory of drive B"), **fixpath()** will return a search path of

b:..*.*

and a prefix, or lead-in, path of

b:..\

However, if the input string was

b:\mydir

it would return a search path of

b:\mydir*.*

and a lead-in path of

b:\mydir\

The search path is for input to functions 4E/4F. If function 4E returns an error, the original path was ungrammatical somehow. Otherwise, the search path will return the same filenames that DIR would print if it were given the same input path you gave to **fixpath()**. When you append one of these returned filename strings to the prefix path set up by **fixpath()**, you have a complete pathname to give to DOS functions 3Dh (open), 41h (erase), 43h (get/set attributes) or 56h (rename).

DDJ

(Listing begins on page 16)

AT™ Pfantasies for your PC or XT.™

Want better speed and memory on your PC or XT without buying an AT?

You've got it!

Phoenix's new Pfaster™286 co-processor board turns your PC or XT into a high-speed engine 60 percent faster than an AT. Three times faster than an XT. It even supports PCs with third-party hard disks. But that's only the beginning.

You can handle spreadsheets and programs you never thought possible. Set up RAM disks in both 8088 and 80286 memory for linkage editor overlays or super-high-speed disk caching. All with Pfaster286's 1mb of standard RAM, expandable to 2mb, and dual-mode design.

You can develop 8086/186/286 software on your XT faster. Execute 95 percent of the application packages that run on the AT, excluding those that require fancy I/O capabilities your PC or XT hardware just isn't designed to handle. Queue multi-copy, multi-format print jobs for spooling. Or, switch to native 8088 mode to handle

hardware-dependent programs and back again without rebooting. All with Pfaster286's compatible ROM software. And, Pfaster286 does the job unintrusively! No motherboard to exchange. No wires to solder. No chips to pull. Just plug it into a standard card slot, and type the magic word, "PFAST."

If you really didn't want an AT in the first place, just what it could do for you, call or write: Phoenix Computer Products Corp., 320 Norwood Park South, Norwood, MA 02062; (800) 344-7200. In Massachusetts, 617-762-5030.

Programmers' Pfantasies™ by

Phoenix

Circle **no. 91** on reader service card.

SEE US AT FALL COMDEX

Dr. Dobb's Clinic Listing   (Text begins on page 10)

```
/* ==============================================================
    fixpath() processes a DOS pathname for two different uses.
The input path, *ip, is presumably a command operand like the
first operand of DIR.  One output, the search path *sp, is
the input possibly augmented with "*.*" or "\*.*" so that it
will work reliably with DOS functions 4E/4F.  The other output
is a lead-in path, *lip, which can be prefixed to the simple
filename.ext returned by functions 4E/4F to make a complete
path for opening or erasing a file.
    The function returns 1 if it is successful, but 0 if the
input path can't be processed and should not be used.
    Some input paths can be processed here yet be invalid or
useless.  The search path produced from such an input will
cause an error return from function 4E (search first match).
=============================================================== */
int fixpath(ip,sp,lip)
register char
        *ip, /* the input path */
        *sp, /* the search path */
        *lip; /* the lead-in or prefix path */
{
char *cp; /* work pointer for scanning paths */
/* ==============================================================
    Pick off 4 special cases:
        (1) the null string, which we take to mean "*.*"
        (2) the simple "d:" reference to a drive, which we
            also take to mean "d:*.*"
        (3) the root-dir reference "d:\" which is mishandled
            by function 4300 of both DOS 2.1 and 3.0.
        (4) any reference that ends in "\"
In all cases, the input path is ok as a lead-in, but it needs
the global qualifier *.* added for searching.
=============================================================== */
        if ( (strlen(ip)==0)                   /* null string */
          ||(strcmp(ip+1,":")==0)              /* d: only */
          ||(ip[strlen(ip)-1]=='\\') )         /* ends in bkslsh */
        {
                strcpy(lip,ip);                /* input is ok prefix */
                strcpy(sp,ip);
                strcat(sp,"*.*");              /* add *.* for search */
                return(1);
        }
/* ==============================================================
    Ok, we have a non-null string not ending in \ and not a lone
drive-letter.  Four possibilities remain:
        (1) an explicit file reference,  b:\mydir\mydat
        (2) an explicit directory reference, \mydir
        (3) an ambiguous file reference, *.* or b:\mydir\*.dat
        (4) an unknown name, a:noway or b:\mydir\nonesuch.dat
We can separate this with fair reliability using DOS function
4300h, get attributes from path.
    The DOS operations used here are based on the MS-C 3.0
library functions, but all C compilers have similar library
routines and header files.
=============================================================== */
        regs.x.ax = 0x4300;     /* AH=43h, AL=00h */
        (char *)regs.x.dx = ip; /* DS:DX -> input path */
        regs.x.cx = 0;          /* clear CX */
        intdos(&regs,&regs);    /* call DOS */

        if ( (regs.x.cflag) && (regs.x.ax==0x0002) )
        {
        /* ==============================================================
            The path is valid, in that all directories and drives
        named in it are valid, but the final name is unknown. No
        files will be found in a search, so quit now.
        =============================================================== */
                return(0);
        }

        if ( ((0!=regs.x.cflag)&&(regs.x.ax==0x0003))
          ||((0==regs.x.cflag)&&(0==(regs.x.cx & 0x0010))) )
        {
```
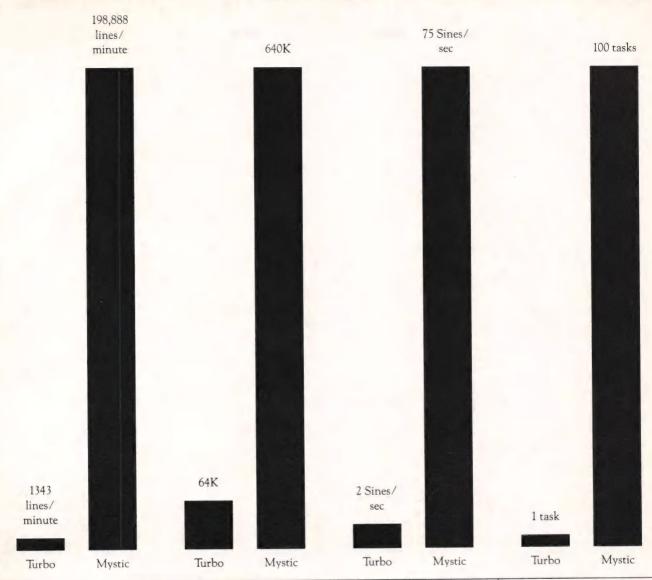
MYSTIC PASCAL vs. TURBO PASCAL

198,888 lines/minute

640K

75 Sines/sec

100 tasks

1343 lines/minute

64K

2 Sines/sec

1 task

| Turbo | Mystic | Turbo | Mystic | Turbo | Mystic | Turbo | Mystic |

COMPILER SPEED

Mystic Pascal uses a completely new compiler design to give extremely fast compilation of ISO Standard Pascal programs. It compiles each line of code the instant you key it into the Full Screen Editor. (Bar Graph is based on compiling a 179 line program with Turbo 3.0 on IBM PC.)

MAXIMUM CODE SIZE

Mystic Pascal lets you use up to 640K for your program's code and data. Your Personal Computer cost thousands of dollars, why settle for a programming language that only lets you use one tenth of its power?

ARITHMETIC SPEED

Mystic Pascal's floating point arithmetic does not use an 8087 but runs almost like it did! We use a special way of storing numbers in memory to achieve single precision arithmetic that is 5 to 50 times faster than other compilers. If you still need 8087 support, we'll be offering that soon.

MULTI-TASKING

The Mystic Pascal system is based on multi-tasking. The Editor, Compiler and other components operate concurrently. Up to 100 of your Pascal procedures may also execute concurrently. And they may communicate by passing messages through queues.

MYSTIC PASCAL FEATURES

- Interactive Pascal
- EXE file output
- 8086 optimized code
- Full Screen Editor
- ISO Standard Pascal
- Help Windows for Pascal Language
- DOS INTeRrupt calls

Mystic Pascal requires an IBM PC or compatible with 256K. Turbo Pascal is a registered trademark of Borland International, Inc.

QUANTITY DISCOUNTS

The price below includes the disk, 75 page manual and shipping within the US and Canada. We can optionally ship by UPS Blue Label (2nd Day Air) within the U.S. Foreign orders add 10% shipping, minimum $15.

Quantity	Price Each	2nd Day Air Price Each
1	$64	$4.00
2- 4	56	3.00
5-10	48	2.50
11-20	42	2.00
21-50	38	1.50

USE YOUR CREDIT CARD TO ORDER NOW BY PHONE (505) 757-6344.

Mystic Canyon Software
P.O. Box 1010
Pecos, New Mexico 87552

Quantity _____
Price _____
Blue Label _____
Total _____

Name _____
Address _____
City _____ State _____ Zip _____

Sorry, CODs and Purchase Orders are NOT accepted. Payment must be in US funds on a US bank.

☐ Check / Money Order   ☐ Visa   ☐ Mastercard
Card _____ Exp. _____
Signature _____

```
/* ======================================================
    Error 3, path not found, could mean total junk or a
misspelt directory name, but most likely it just means
the path ends in an ambiguous name.  If there's an error
the initial search (function 4Eh) will fail.
    With no error and reg cx NOT showing the directory
attribute flags 0010, we have a normal, unambiguous file
pathname like "b:\mydir\mydat" or just "myprog.com".
    In either case the search path is the whole input
path.  The lead-in is that shorn of whatever follows the
rightmost \ or :, whichever comes last -- or just a null
string if there is no \ or :.
====================================================== */
        strcpy(sp,ip); /* working copy of ip in sp */
        cp = sp+strlen(sp)-1;
        for(; cp > sp; --cp)
                if (('\\'==*cp)||(':'==*cp)) break;
        if (cp>sp) ++cp; /* retain colon/slash */
        *cp='\0'; /* may make a null string */
        strcpy(lip,sp);
        strcpy(sp,ip); /* whole input for search */
        return(1);
}

if ((0==regs.x.cflag)&&(regs.x.cx & 0x0010))
{
/* ======================================================
    No error and the directory attribute returned in cx
shows an unambiguous path that happens to end in the
name of a directory, for instance "..\..\bin" or
"b:\mydir"  Applying the rules of DIR or COPY, we
have to treat these as global refs to all files named
in the directory.  The search path is the input with

"\*.*" tacked onto it.  The lead-in path is just the
input plus a backslash.
====================================================== */
        strcpy(sp,ip);
        strcpy(lip,ip);
        strcat(sp,"\\*.*");
        strcat(lip,"\\");
        return(1);
}
else
{ /* unexpected events make me nervous, give up */
        return(0);
}
}
```

End Listing

If lightning still scares you, you're using the wrong file manager.

Be sure. Btrieve™.

Lightning may strike. But it doesn't have to destroy your database.

Btrieve™ file management offers automatic file recovery after a system crash. So accidents and power failures don't turn into database disasters. Your Btrieve-based applications will come up when the lights come back on.

Fast. Btrieve is lightning fast, too. Its b-tree file structure automatically balances—you never waste time reorganizing the index. And Btrieve is written in Assembly language especially for the IBM PC™. The result: electrifying speed on file maintenance routines. Applications that run fast. Users who don't waste time waiting.

The standard for networking. When your application requires multi-user file sharing, Btrieve/N (network version) sets the standard for the industry's most popular LANs: IBM's PC Network, Netware™, Davong MultiLink™, Omninet™, PC Net™, EtherSeries™, Nestar™, and NetOne™. Btrieve/N offers **safe** network file management that coordinates simultaneous updates and protects against lost data.

Fully-relational data management. Using SoftCraft's entire family of products gives you a complete, fully-relational database management system. Rtrieve™ adds report writing capabilities for generating the reports you need. Xtrieve™ speeds users through database queries with interactive, on-screen menus—no command language or special syntax.

For professional programmers. Btrieve is the fast, reliable answer for all your application development. In any development language—BASIC, Pascal, Cobol, C, Fortran, and APL. With multikey access to records. Unlimited records per file. Duplicate, modifiable, and segmented keys. Variable cache buffer.

With Btrieve, you can develop better applications faster. And know they'll be safe if lightning strikes.

Suggested retail prices: Btrieve, $245; Btrieve/N, $595; Xtrieve, $195; Xtrieve/N, $395; Rtrieve, $85; Rtrieve/N, $175. Requires PC-DOS or MS™-DOS 1.X, 2.X, or 3.X.

Btrieve, Xtrieve, and Rtrieve; IBM; Netware; Davong MultiLink; Omninet; PC Net; EtherSeries; Nestar; NetOne; and MS are trademarks of SoftCraft Inc.; International Business Machines; Novell Data Systems; Davong Systems Inc.; Corvus Systems; Orchid Technology; 3Com Corp.; Nestar Systems Inc.; Ungermann-Bass; and Microsoft Inc.

SoftCraft Inc.

P.O. Box 9802 #917
Austin, Texas 78766
(512) 346-8380   Telex 358 200

C CHEST

by Allen Holub

This month we're going to look at an implementation of Knuth's hyphenation algorithm. This code is yet another part of my version of the Unix text formatter, nroff.

Bugs and Fixes

Before getting started, here's a little miscellany. First, while bringing up a version of **printf**() the other day, I found a bug in Version 2.15 of Lattice C (though the developer would probably argue that it's a feature).

When writing a subroutine with a varying number of arguments of varying types (like **printf**), arguments are fetched from the stack frame by casting a "generic" pointer to point at an individual object on the stack. After fetching the argument, the pointer is advanced past it. In other words, we want to cast a character pointer into a long pointer, fetch a long object off the stack, and then advance the character pointer by the size of a long. The first thing I tried was:

```
char *p ;
long x ;

x = *(long *)p++ ;
```

but this won't work correctly. Here, the problem is order of precedence. *****, **(long *)**, and **++** are all at the same precedence level, but they associate right to left. When fully parenthesized, the associativity gives us:

```
x = *((long *)(p++))
```

Because the **++** binds more tightly to **p** than the cast, the compiler will assume that **p** is a character pointer (rather than a long pointer) for the sake of the increment, though it will fetch the long object correctly. In other words, the expression evaluates

to a long, but 1 (rather than 4) is added to the pointer. I tried to force the precedence to behave by doing:

```
x = *( (long *)p )++ ;
```

but now Lattice gives me an error 25 (lvalue required) and won't compile it. By itself, **((long *)p)++** doesn't compile either. Because this statement indeed has an lvalue (the **x**), I was rather nonplused. I think that the compiler was confused by the cast, thinking that I was trying to do something such as **x++ = y** or:

```
int foo;
(long)foo = 5L;
```

which won't work because this would force **foo** to overflow (only an integer's worth of space is reserved for **foo**, and we're trying to squeeze a long object into it). However, **((long *)p)++** should work because we're just modifying the way that the pointer arithmetic works, not overflowing the target. A more clever error detection algorithm would take care of this exception. The following, though inelegant, does work:

```
char *p;

x = *(long *)p ;
p += sizeof( long );
```

While we're on the subject of bugs, several people have written about what we thought was a bug in **ls**, the directory utility printed a couple months back. The problem here is underlining. On a Compaq, directory names comes out at half intensity instead of underlined. The problem here is the hardware (and I contend that this is a bug, not a feature). The Compaq emulates an IBM color card running a monochrome monitor. The

color card doesn't support underlining (I think that it should) but prints text in a different color rather than underlining it. On a monochrome display, this second color comes out as reduced intensity. All the IBM cards, color or otherwise, should support the attributes specified in the ANSI.SYS documentation. Anyway, the Paradise Modular Graphics card has the same problem as the Compaq. On the other hand, the Hercules Graphics card does underlining correctly.

There's a second bug in **ls** that's a little more serious. **Ls**, as it stands, runs fine under DOS Version 3.0. However, when running under Version 2.x, it couldn't find the root directory (\) when it was explicitly requested by an **ls ** command. A version-independent fix for the problem is given in Listing One (page 26).

Hyphenation

Moving on, the capability to hyphenate words is desirable in any text formatter or word processor. A lot of needless white space between words can be eliminated by hyphenation. Narrow columns look better, and we can get more words per page. Looking at the output of several commercially available word processors, I suspect that most of them derive their algorithms from the one given by Knuth in T_EX *and Metafont*, and I've used his algorithm here.[1]

Hyphenation is a harder problem than you would think at first. To use Knuth's examples, why is *record* hyphenated *rec-ord* when used as a noun and *re-cord* when as a verb? Consequently, it's impractical to attempt a truly universal algorithm—there are just too many exceptions. Knuth's algorithm is intentionally conservative—if it isn't absolutely sure how to split two syllables, it doesn't split them.

Algorithm Description

Knuth attacks the problem by breaking it into several parts, the first of which is exception removal. He keeps a dictionary of words that are known exceptions to his other rules and then sees if a word is an exception before processing it further. Strictly speaking, we could use a huge dictionary and dispense with the rest of the algorithm, but this would be too cumbersome to be practical. Knuth includes a dictionary of about 300 words in his book, some of which are rather useless. (How many times do you use "coneflower" in an average document?) There just wasn't room to print the dictionary look-up part of the hyphenator this month. It's a straightforward program—it reads a sorted list of exceptions and then does a simple binary search in a table.

The second part of the algorithm does suffix removal. The following suffixes are recognized:

> -able, -ary, -cal, -cate, -cial, -cious, -cient, -dent, -ful, -ize, -late, -less, -ly, -ment, -ness, -nary, -ogy, -rapher, -raphy, -scious, -scope, -scopic, -sion, -sphere, -tal, -tial, -tion, -tional, -tive, and -ture.

Some of these are qualified

> -able must be preceded by e, h, i, k, l, o, u, v, w, x, y, nt, or rt; -ary must be preceded by ion or en; -cate and -late must be preceded by a vowel (a, e, i, o, u, or y); -nary may not be preceded by e or io; -cious may not be preceded by s; and -ize must be preceded by l.

Another complication is -ing. Knuth says if -ing is preceded by fewer than four letters, don't insert a hyphen. If -ing is preceded by two identical consonants (other than f, l, s, or z) put the hyphen between the consonants, and if it's preceded by any letter but l, break as -ing. If the letter before -ling is b, c, d, f, g, k, p, t, or z, break before this letter (but do not break a ck combination before -ling);

otherwise break -ing.

If an -able, -ary, -ful, -ize, -less, -ly, -ment, or -ness is recognized, then the algorithm is applied again to the residual word (the word less the suffix we just found). Similarly, if we fail to find a suffix and the word ends in s, e, or ed, we try again with the s, e, or ed removed.

Having removed suffixes, we now try to remove potential prefixes, working on the word without all the suffixes just recognized. The following prefixes will be recognized:

> be-, com-, con-, dis-, equi-, equiv-, ex-, hand-, horse-, hy-per-, im-, in-, in-ter, in-tro-, lex-i-, mac-ro-, math-e-, max-i-, min-i-, mul-ti-, non-, out-, over-, pseu-do-, quad-, semi-, some-, sub-, su-per-, there-, trans-, tri-, un-der, and un-.

The intermediate hyphens aren't inserted unless the entire prefix is found (e.g., hy-per-bole vs. hyp-no-tic). There are also exceptions here:

Be- must be followed by *c, h, s,* or *w; dis-* may not be followed by *h* or *y; equi-* may not be followed by *v; trans-* must be followed by *a, f, g, l,* or *m; tri-* must be followed by *a, f,* or *u; un-* may not be followed by *der* or *i.*

As with suffixes, the algorithm is applied repeatedly after *dis-, im-, in-, non-, over-,* or *un-* is recognized.

At this juncture, the word has been partitioned into three parts: prefix, root, and suffix. We now apply several consonant-pair rules to the root portion of the word. Here, vowels are *a, e, i, o, u,* and *y,* and consonants are all other single letters. Moreover, the pairs *ch, gh, ph, sh,* and *th* are treated as single consonants. If a vowel-consonant-consonant (VCC) pattern is found and both consonants are the same and neither consonant is an *l* or *s,* the break will be between the consonants. In the case of a *vowel-ll* or *vowel-ss* pattern, there will be a break only if the following letter isn't a vowel and the word doesn't end VCC*er* or VCC*ers.* If the pattern is *vowel-ck,* the break will be after the *ck.* If the pattern is *vowel-qu,* the break will be before the *qu.*

If a VCCV combination is found, the break will be between the consonants unless they are one of the following combinations:

> *bl, br, cl, cr, chl, chr, dg, dr, fl, fr, ght, gl, gr, kn, lk, lq, nch, nk, nx, phr, pl, pr, rk, sp, sq, tch, tr, thr, wh, wl, wn,* or *wr.*

Finally, don't break between consonants if the word ends with VCC*er,* VCC*ers,* VCC*age,* VCC*ages,* VCC*est* and the two consonants are either *ft, ld, mp, nd, ng, ns, nt, rg, rm, rn, rt,* or *st.*

Once the word is hyphenated, Knuth goes back and fixes obvious mistakes. He takes back very short syllables and syllables ending with a silent *e.* I decided to use the exception dictionary to do this rather than adding more code to the algorithm prop-er. I should say that Knuth is not his usual lucid self when describing this part of the algorithm. A major factor in my not implementing the "short end" retention is that I can't, for the life of me, figure out what he was saying. If anybody can decipher this part of the algorithm description, please enlighten me.

Implementation

The code this month (Listing Two, page 26) supports the suffix removal, prefix removal, and consonant-pairs part of the algorithm. The Lattice compiler produces a 9063-byte object module (which seems a little bloated to me, though the code could be optimized a bit more if space is a problem). Hyphens are marked by setting the high bit of the character that will follow the hyphen (the first letter in the syllable). Only unhyphenated words composed of lower case letters will be processed. Among other things, this lets us avoid the problems associated with proper nouns, which shouldn't be hyphenated.

All global variables and all subroutines that don't need to be accessed from another module are declared static. The one externally accessible routine is the driver, **hyphen( )** (on line 484), which calls the various other processing routines.

The suffix/prefix removal routines use table-driven state machines for pattern recognition. As usual, there is a trade-off between space and execution time. The fastest state machines use a large two-dimensional array in which one axis is the current state and the other the current input character. The table itself holds the next state. As most of the states have only one legal transition, this method uses an unnecessarily large amount of memory. At the other extreme, we can keep a list associated with each state that contains all legal input characters and the state to go to when that character is encountered. This method usually takes less space, but we have to search the list every time we process a character so execution time isn't great when the lists are too long.

The tables used here are a compromise between these two methods. Each table is an array of pointers to

Prefixes:

0	Failure, return original pointer
1-68	Get another character and branch to next state as indicated in table.
69	Get a character, hyphenate *p, branch to next state using *p as the current character.
70	hyphenate *(p−1) and branch as above.
71-80	don't exist.
81	hyphenate *(p−1) and return (p−1).
82	hyphenate *p and *(p−1), return p.
83	hyphenate *p and *(p−2), return p.
84	hyphenate *p and *(p−3), return p.
85	hyphenate *(--p), don't return.
86	hyphenate *p, don't return.
87	hyphenate *p and return p.

Suffixes:

0	Failure, if the word ends in *s, e* or *ed* strip it off and try again, else return original string pointer.
1-57	Get another character and branch to next state as indicated in table.
58-79	Don't exist.
80	p += 1, hyphenate *(p + 1).
81	p += 2, hyphenate *(p + 1).
82	p += 3, hyphenate *(p + 1).
83	hyphenate *(p + 1).
84	Failure, return original pointer.
85	Doesn't exist.
86	process an -ing.
87	hyphenate *(p + 1) and return p.
88	hyphenate *(p + 2) and return (p + 1).
89	hyphenate *(p + 5) and *(p + 1), return p.

In both tables, unspecified transitions go to state 0. Unmarked transitions always go to the indicated state, regardless of input character. **P** points at the character being processed at present.

Figure 1
States and Associated Actions

character arrays. The tables are indexed by current state. If there's only one legal transition out of a given state, the first byte of the associated array is 1, the next byte is the one legal input character, and the third byte is the state to go to when that character is found. All other input characters will cause a transition to state 0. The second type of array is used when there is more than one legal transition out of a state. In this case, the first byte of the array is 0 and the next 26 bytes are an array of next states indexed by current input character. The state tables begin on line 605 of the listing.

The state diagrams for these tables are shown in Figure 2 (page 24) and Figure 3 (page 25). In both machines, state 0 is a failure state. States with numbers greater than or equal to 80 are success states. These are the only states in which any sort of action is performed. The actions themselves are spelled out in Figure 1 (page 22). Note that the suffix recognition code goes through the word from back to front. Consequently, the

state table for suffixes is a little weird looking (all the suffixes are backwards).

The table-processing routines start on lines 110 (suffix) and 221 (prefix). They are essentially the same, though different tables are used to do the processing. The subroutine next( ), when given the current state and input character, will return the next state. All possible action states are taken care of in the switches on lines 155 and 235.

Consonant-pair recognition is done by brute force rather than with a machine. (I tried to do a state machine, but it got too complicated). One of the places where the code size could be reduced is here. Nextch( ) (on line 318) gets the next input character and advances the string pointer past the character. If a consonant pair (*th*, *sh*, *ph*, *ch*, or *gh*) is found, both characters are skipped and an integer value (#**defined** on line 31 f.) is returned. The actual consonant-pair processing routine starts on line 393.

That's hyphenation. The algorithm works correctly with most of the

words I've tried it on. My version successfully hyphenates su-per-califragi-lis-ticex-pialido-cious, just like Knuth's does. (Note that the algorithm doesn't insert every possible hyphen, but it doesn't put in any incorrect ones either.)

If anyone doesn't feel like typing in this month's program, machine readable copies of all source code from C Chest are now available on 5¼-inch IBM PC compatible disks from Software Engineering Consultants, P.O. Box 5679, Berkeley, CA 94705. The cost is $25 per disk (one disk is one month's column). In addition, my version of grep, the Unix generalized regular expression parser presented in *Doctor Dobb's Journal* #96, is still available for $35 from the same address.

Notes

[1] Donald E. Knuth, T_EX and Meta-font , *New Directions in Typesetting* (Digital Press),pp.180–186. DDJ

(Listings begin on page 26)

Reader Ballot
Vote for your favorite feature/article.
Circle Reader Service **No. 191**.

Figure 1 SUFFIXES

Figure 2 PREFIXES

C Chest (Text begins on page 20)
Listing One

```
+---------------------------------------------------------------+
|          Changes to ls.c to fix the root directory problem.   |
|                                                               |
|          1) Add the subroutine isrootdir():                   |
+---------------------------------------------------------------+

isrootdir( name )
char *name;
{
        if( *name && name[1] == ':' )
                name += 2;

        return( (*name == '\\' || *name == '/') && !name[1] );
}

+---------------------------------------------------------------+
| 2) Modify the first if statement of the subroutine fixup_name()|
|              (line 184 of the original listing) to read:      |
+---------------------------------------------------------------+

        if( !isrootdir(name)  && !find_first(name, ALL, regs) )
```

End Listing One

Listing Two

```
1: #include <stdio.h>
2: #define DEBUG
3:
4: /*-----------------------------------------------------------------*
5:  *       HYPHEN.C - An implementation of the Knuth hyphenation algorithm  *
6:  *                                                                 *
7:  *            Copyright (c) 1985, Allen I. Holub.  All rights reserved.   *
8:  *-----------------------------------------------------------------*
9:  */
10:
11: /* Various psuedo-subroutines. HYPHEN defines a bit to set when a hyphen
12:  * is inserted. HYPHENATE sets the bit, UNHYPHENATE clears it,
13:  * HAS_HYPHEN tests for it. The ER macro checks for an "er" at the end
14:  * of a word, it's used by the consonant pair checking routine.
15:  * Is consonant returns true if c is a consonant.
16:  */
17:
18: #define HYPHEN          0x80
19: #define HYPHENATE(c)    ( (c) |= HYPHEN )
20: #define UNHYPHENATE(c)  ( (c) &= ~HYPHEN )
21: #define HAS_HYPHEN(c)   ( (c) &   HYPHEN )
22:
23: #define ER(p,end)       (*p == 'e' && *(p+1) == 'r' &&  (p+1) == end)
24: #define isconsonant(c)  ((c) && !isvowel(c))
25:
26: /*      The dipthongs ch, gh, ph, sh and th are treated as single
27:  *      consonants. The subroutine nextch() will map these two
28:  *      characters into a single character as follows:
29:  */
30:
31: #define CH ('z' + 1 )   /* {      0x7b  \173    */
32: #define GH ('z' + 2 )   /* |      0x7c  \174    */
33: #define PH ('z' + 3 )   /* }      0x7d  \175    */
34: #define SH ('z' + 4 )   /* ~      0x7e  \176    */
35: #define TH ('z' + 5 )   /* DEL    0x7f  \177    */
36:
37: /*      S(x) is used to initialize the state tables. It turned out to
38:  *      be easier to use a define than a typedef. ATYPE is used store
39:  *      strings in something other than chars (useful if we're keeping
40:  *      attributes around). I haven't tried compiling with ATYPE set to
41:  *      anything except char so be careful if you do.
42:  */
43:
44: #define S(x)  static char x[]
45: typedef char   ATYPE;
46:
47: #define LATTICE 1          /* Compile for Lattice C compiler      */
48:
49: /*-----------------------------------------------------------------
50:  *      GLOBAL VARIABLES:
51:  */
52:
53: static char **States;      /* Points to table for current state machine */
54: extern char *Suffixes[];   /* Suffix state table (at end of this module) */
55: extern char *Prefixes[];   /* Prefix state table          "             */
56:
57: #ifdef DEBUG
58: static int  Debug = 0;   /* True if debug diagnostics are to be printed */
59: #endif
60:
61: /*-----------------------------------------------------------------*/
62:
63: #ifdef LATTICE
64:
```

C Chest (Listing continued, text begins on page 20)
Listing Two

```
65:  extern   char      *stpchr();
66:  #define index(s,c)   stpchr(s,c) /* For some reason the unix index()
67:                                    * function is called stpchr() by
68:                                    * Lattice.
69:                                    */
70:  #define register     static      /* This will give more efficient code
71:                                    * since Lattice doesn't support any
72:                                    * register variables. DON'T EXPLICITLY
73:                                    * INITIALIZE AUTOMATIC VARIABLES IF
74:                                    * THIS #DEFINE IS ACTIVE.
75:                                    */
76:
77:  #else      /*- - - - - - - - - - - - - - - - - - - - - - - - - - - */
78:
79:  char     *index(s,c)
80:  char     *s;
81:  {
82:           /*      Return a pointer to the left-most occurance of the
83:            *      character c in the string s or NULL if the character
84:            *      isn't found.
85:            */
86:
87:           for(; *s ; s++ )
88:                   if( *s == c )
89:                           return s;
90:           return NULL;
91:  }
92:
93:  #endif
94:
95:  /*-----------------------------------------------------------------------*/
96:
97:  isvowel(c)
98:  ATYPE   c ;
99:  {
100:          /* We check for a vowel in the order in which it is likely to
101:           * appear in English (ETAOIN SHRDLU). Return true if c is a vowel.
102:           */
103:
104:          c &= 0x7f;
105:          return( c=='e' || c=='a' || c=='o' || c=='i' || c=='u' || c=='y');
106:  }
107:
108:  /*-----------------------------------------------------------------------*/
109:
110:  static  ATYPE   *suffix( beg, end )
111:  ATYPE   *beg, *end ;
112:  {
113:          /*      Split off any suffixes, using the "Suffixes" state table
114:           *      (defined below) to recognize them.
115:           */
116:
117:          register char   *p;
118:          register int    state, c, c2, times ;
119:
120:          state  = 1;
121:          times  = 0;
122:          States = Suffixes;
123:
124:          for( p = end; p >= beg ;)
125:          {
126:                  state = next( state, c = *p-- & 0x7f );
127:
128:                  if( !('a' <= c && c <= 'z') ) /* Words containing non-
129:                          return end;           /* lower-case characters
130:                                                /* won't be hyphenated.  */
131:
132:                  /* Test for state 0 with an if instead of in the switch.
133:                   * Lattice C will make the switch very inefficient if the
134:                   * range of case values is too high. State 0 is wierd;
135:                   * we use it to test for trailing e or ed on failure.
136:                   * The first time through we set p to end so we can check
137:                   * the end of the word. The third time through we abort.
138:                   * This means that we'll retry if the word ends -s -e -d
139:                   * -ed -de -dd -ee. The extra endings shouldn't matter.
140:                   */
141:
142:                  if( state == 0 )
143:                  {
144:                          if( times == 0 )
145:                                  p = end ;       /* Strip trailing e or d */
146:
147:                          else if( times == 1 ) /* Strip trailing ed, de */
148:                                  p = end-1 ;   /* ee, or dd            */
149:                          else
150:                                  return end;
151:
152:                          times++;
153:                  }
154:
155:                  switch( state )
156:                  {
```

C Chest (Listing continued, text begins on page 20)
Listing Two

```
157:                         case 86:
158:                                 /*
159:                                  * Process -ing. This could have been done with
160:                                  * more states but it seemed easier to do it here.
161:                                  */
162:
163:                                 c  =  *p    & 0x7f ;
164:                                 c2 = *(p-1) & 0x7f ;
165:
166:                                 if( p - beg < 3 )
167:                                 {
168:                                         return end;     /* FAIL */
169:                                 }
170:                                 else if(     ( c == c2 ) && ( !isvowel(c) )
171:                                         && ( c != 'f') && ( c != 's'   )
172:                                         && ( c != 'l') && ( c != 'z'   ) )
173:                                 {
174:                                         --p;
175:                                 }
176:                                 else if( c == 'l'  &&  index("bcdfghkptz", c2)  )
177:                                 {
178:                                         p -= (c2 == 'k' && *(p-2) & 0x7f)) == 'c')
179:                                                                 ? 1 : 2 ;
180:                                 }
181:
182:                                 /* fall through to state 87 */
183:
184:                         case 87:
185:                                 HYPHENATE( *(p+1) );
186:                                 return p;
187:
188:                         case 88:
189:                                 HYPHENATE( *(p+2) );
190:                                 return( p + 1 );
191:
192:                         case 89:
193:                                 HYPHENATE( *(p+5) );
194:                                 HYPHENATE( *(p+1) );
195:                                 return p;
196:
197:                         case 82: p++;   /* p += 3 */
198:                         case 81: p++;   /* p += 2 */
199:                         case 80: p++;   /* P += 1 */
200:                         case 83:
201:                                 end = p ;
202:                                 HYPHENATE( *(p + 1) );
203:                                 state = 1;
204:                                 break;
205:
206:                         case 84:                        /* FAILURE STATE */
207:                                 return end;
208:                 }
209:         }
210:
211: #ifdef DEBUG
212:         if( Debug )
213:                 printf("\n");
214: #endif
215:
216:         return end;
217: }
218:
219: /*-----------------------------------------------------------------------*/
220:
221: static  ATYPE  *prefix( beg, end )
222: ATYPE  *beg, *end;
223: {
224:         /*      Split off any prefixes, using the "Prefixes" state table
225:          */
226:
227:         register ATYPE  *p;
228:         register int    state, c ;
229:
230:         state  = 1;
231:         States = Prefixes;
232:
233:         for( p = beg ; p < end ;)
234:         {
235:                 switch( state = next(state, c = *p++ & 0x7f) )
236:                 {
237:                 case 82: HYPHENATE( *p ); HYPHENATE( *(p-1) ) ; return p;
238:                 case 83: HYPHENATE( *p ); HYPHENATE( *(p-2) ) ; return
239:                 case 84: HYPHENATE( *p ); HYPHENATE( *(p-3) ) ; return p
240:
241:                 case 81: --p;
242:                 case 87: HYPHENATE( *p );
243:                         return p;
244:
245:                 case 85: --p;
246:                 case 86: beg  = p;
247:                         state = 1;
248:                 case 69: HYPHENATE( *p );
```

```
249:                          break;
250:
251:                  case 70: HYPHENATE( *(p-1) );
252:                          break;
253:
254:                  /*      The following states don't actually exist.
255:                   *      Putting them in the table will cause Lattice to
256:                   *      compile the switch in a more efficient manner.
257:                   *      (see p 4-33 of the 1985 vintage manual).
258:                   */
259:
260:                  case 71: case 72: case 73: case 74: case 75:
261:                  case 76: case 77: case 78: case 79:
262:                          break;
263:          }
264:      }
265:
266: #ifdef DEBUG
267:          if( Debug )
268:                  printf("\n");
269: #endif
270:
271:          return beg;      /* Failure */
272: }
273:
274: /*-------------------------------------------------------------------
275:  *      This table is used to find exceptions to the VCCV rule. It is
276:  *      indexed by the first of the two consonents & contains pointers
277:  *      to strings, any character of which will form an exception if
278:  *      it's the second consonant.
279:  */
280:
281: static char      *vccv_except[] =
282: {
283:          /* a  */          ""        ,
284:          /* b  */          "lr"      ,
285:          /* c  */          "lr"      ,
286:          /* d  */          "gr"      ,
287:          /* e  */          ""        ,
288:          /* f  */          "lr"      ,
289:          /* g  */          "lr"      ,
290:          /* h  */          ""        ,
291:          /* i  */          ""        ,
292:          /* j  */          ""        ,
293:          /* k  */          "n"       ,
294:          /* l  */          "kq"      ,
295:          /* m  */          ""        ,
296:          /* n  */          "{kx"     , /* CH, k, x */
297:          /* o  */          ""        ,
298:          /* p  */          "lr"      ,
299:          /* q  */          ""        ,
300:          /* r  */          "k"       ,
301:          /* s  */          "pq"      , /* CH, r    */
302:          /* t  */          "{r"      , /* CH, r    */
303:          /* u  */          ""        ,
304:          /* v  */          ""        ,
305:          /* w  */          "hlnr"    ,
306:          /* x  */          ""        ,
307:          /* y  */          ""        ,
308:          /* z  */          ""        ,
309:          /* CH */          "lr"      ,
310:          /* GH */          "t"       ,
311:          /* PH */          "r"       ,
312:          /* SH */          ""        ,
313:          /* TH */          "r"       ,
314: };
315:
316: /*----------------------------------------------------------------------*/
317:
318: static ATYPE    nextch( pp, endp )
319: ATYPE    **pp, *endp;
320: {
321:          /* Advance *pp (one for a vowel or consonant, 2 for a dipthong)
322:           * and return the character we've just skipped. Dipthongs are
323:           * mapped to the single characters: TH, SH, PH, CH or GH.
324:           * 0 is returned once *pp reaches or passes endp.
325:           */
326:
327:          register ATYPE  rval, *p;
328:
329:          if( (p = *pp)  > endp )
330:                  return (ATYPE) 0;
331:
332:          rval = *p++ & 0x7f ;
333:
334:          if( ( *p & 0x7f )  ==  'h' )
335:          {
336:                  switch( rval )
337:                  {
338:                  case 't':        rval = TH;      p++;    break;
339:                  case 's':        rval = SH;      p++;    break;
340:                  case 'p':        rval = PH;      p++;    break;
341:                  case 'c':        rval = CH;      p++;    break;
342:                  case 'g':        rval = GH;      p++;    break;
343:                  }
```

(Continued on next page)

```
344:                }
345:
346:     #ifdef DEBUG
347:            if( Debug )
348:                    printf( "nextch(), got <%c>\n", rval );
349:     #endif
350:
351:            *pp = p;
352:            return rval;
353:     }
354:
355:
356:     /*-------------------------------------------------------------------*/
357:
358:     static   iswierd( x, y, p )
359:     ATYPE    *p;
360:     {
361:            /* Return true if the string pointed to by p ends in:
362:             *      XYer XYers XYage XYages XYest
363:             *
364:             * where XY is one of the pairs:
365:             *      ft ld mp nd ng ns nt rg rm rn rt st
366:             */
367:
368:            register int    c1, c2, c3;
369:
370:            c1 = *p++ ;
371:            c2 = *p++ ;
372:            c3 = *p   ;
373:
374:            return(
375:                    (
376:                            ( c1 == 'e' && c2 == 'r'                )
377:                       ||   ( c1 == 'a' && c2 == 'g'  &&  c3 == 'e' )
378:                       ||   ( c1 == 'e' && c2 == 's'  &&  c3 == 't' )
379:                    )
380:                    &&
381:                    (       ( x == 'f'  && y == 't' )
382:                       ||   ( x == 'l'  && y == 'd' )
383:                       ||   ( x == 'm'  && y == 'p' )
384:                       ||   ( x == 's'  && y == 't' )
385:                       ||   ( x == 'n'  && index( "dgst", y ) )
386:                       ||   ( x == 'r'  && index( "gmnt", y ) )
387:                    )
388:                  );
389:     }
390:
391:     /*-------------------------------------------------------------------*/
392:
393:     static   consonants( beg, end )
394:     ATYPE    *beg, *end;
395:     {
396:            /*        Hyphenate consonant pairs.
397:             *        Look for a VCC pattern: use the state machine:
398:             *
399:             *                    V           C         C
400:             *    Start ----> 1 -----> 2 -----> 3 -----> 4
401:             *            ^    |C   ^    |V       |V      | (no fetch )
402:             *            |    |    |    v        |       |
403:             *            |<--+    +------<-----+        |
404:             *            |                              |
405:             *            +------------------------------+
406:             *
407:             *        The machine is implemente explicitly (with whiles and ifs
408:             *        ifs and such) rather than with a table driven engine.
409:             *        A new character is fetched in states 1, 2, & 3 (but not 4).
410:             *        When you exit reach state 4 c1 and c2 will hold the two
411:             *        consonants, cp will point to the beginning of the second
412:             *        consonant, beg will point just past the second consonant.
413:             */
414:            register ATYPE  c1, c2, *cp;
415:            register char   *p;
416:
417:            while( 1 )
418:            {
419:                    state1:
420:
421:                    do {
422:                            c2 = nextch( &beg, end );
423:                    } while( isconsonant(c2) );
424:
425:                    do {
426:                            for( c1 = c2; isvowel(c1) ; )
427:                            {
428:                                    cp = beg;
429:                                    c1 = nextch( &beg, end );
430:                            }
431:
432:
433:                            if(c1 == 'q' && *beg == 'u')    /*         Vqu */
434:                            {
435:                                    HYPHENATE( *cp );       /*         V-qu */
```

```
436:                                          nextch( &beg, end );      /*  skip the u */
437:                                          goto statel;
438:                          }
439:
440:                          cp = beg ;
441:                          c2 = nextch(&beg, end) ;
442:
443:                  } while( isvowel(c2) ) ;
444:
445:                  if( !c1 || !c2 )
446:                          break;
447:
448:                  /*      At this point we have found a Vowel-Cons-Cons
449:                   *      sequence. C1 and c2 will hold the left and right
450:                   *      consonant respectively. Beg will point at the
451:                   *      character following the second consonant. Cp
452:                   *      will point at the second consant.
453:                   */
454:
455:                  if( c1 == 'c' && c2 == 'k' )            /*          Vck  */
456:                  {
457:                          if( *beg )
458:                                  HYPHENATE( *beg );      /*          Vck- */
459:                  }
460:                  else if( c1 == c2 )
461:                  {
462:                          if( (c1 != 'l' && c1 != 's') ||
463:                              (isvowel(*beg) && !ER(beg,end)) )
464:                                  HYPHENATE( *cp );
465:                  }
466:
467:                  else if( isvowel( *beg ) )              /*          VCCV */
468:                  {
469:                          if( ! iswierd(c1, c2, beg) )
470:                          {
471:                                  if( !((p = vccv_except[ (int)c1 - 'a' ])
472:                                                          && index(p, c2) ))
473:                                          HYPHENATE( *cp );
474:                          }
475:                  }
476:          }
477:
478:          if( HAS_HYPHEN(*end) )
479:                  UNHYPHENATE( *end );
480: }
481:
482: /*----------------------------------------------------------------------*/
483:
484: hyphen( beg, end )
485: ATYPE  *beg, *end;
486: {
487:          /*      Hyphenate the word deliniated by beg and end: First
488:           *      strip suffixes, then strip a trailing s e or ed,
489:           *      then strip prefixes. Only words with more than four
490:           *      letters and which consist of lower case letters only
491:           *      (no NULLs) will be hyphenated. 0 is returned if no
492:           *      attempt was made to hyphenate the word, one otherwise.
493:           *      If the word has already been hyphenated, 1 is returned
494:           *      but the word is not modified.
495:           */
496:
497:          register ATYPE  *prefixp, *suffixp;
498:
499:          if( end-beg <= 4 )
500:                  return 0;
501:
502:          for( prefixp = beg; prefixp <= end ; prefixp++ )
503:          {
504:                  if( HAS_HYPHEN(*prefixp) )
505:                          return 1;
506:
507:                  if( !islower(*prefixp & 0x7f) )
508:                          return 0;
509:          }
510:
511:
512:          suffixp = suffix(beg,end);             /* Hyphenate and remove any   */
513:                                                 /*         suffixes           */
514:          if( suffixp == end )
515:          {
516:                  /*      If root has a trailing s e or ed remove it
517:                   *      before applying any other rules
518:                   */
519:
520:                  if(*end == 's' || *end =='e')
521:                          suffixp = end - 1 ;
522:
523:                  else if( *(end-1) == 'e'  &&  *end == 'd' )
524:                          suffixp = end - 2 ;
525:          }
526:
527:          prefixp = prefix( beg, suffixp);     /* Hyphenate and remove any */
528:                                               /* prefixes                 */
529: #ifdef DEBUG
530:          /*----------------------------------------------------------------*/
```

```
531:              /*   Print the word in the form: prefixes/root/suffixes */
532:              /*                                                      */
533:              bprint( beg, prefixp-1 );       /* Print prefixes     */
534:              putchar('/');
535:              bprint( prefixp, suffixp );     /* Print middle       */
536:              putchar('/');
537:              bprint( suffixp+1, end );       /* Print suffixes     */
538:              /*------------------------------------------------------*/
539: #endif
540:
541:          if( (suffixp - prefixp)  >= 3 )
542:          {
543:                      /* Apply the consonant pairs rules only if the word
544:                       * has at least 4 letters in it (after prefixes
545:                       * and suffixes have been removed
546:                       */
547:
548:                      consonants( prefixp, suffixp );
549:          }
550:
551:          return 1;
552: }
553:
554: /*--------------------------------------------------------------*/
555:
556: static  next( cur_state, cur_char )
557: ATYPE   cur_char  ;
558: int     cur_state ;
559: {
560:          /*      Given the current state and the current input character
561:           *      return the next state. The global variable States must
562:           *      point at the correct state table (Suffixes or Prefixes).
563:           */
564:
565:          char    *p = States[ cur_state ] ;
566:
567: #ifndef DEBUG
568:
569:          if( *p )
570:                      return (int)( (cur_char == p[1]) ? p[2] : 0 );
571:          else
572:                      return (int)( p[(cur_char - 'a') + 1] );
573: #else
574:
575:          int     rval;
576:          if( Debug )
577:                      printf("%s: Current state = %d, Input char = <%c>, ",
578:                                  States == Prefixes ? "PREFIX" : "SUFFIX",
579:                                  cur_state, cur_char );
580:
581:          if( *p )
582:                      rval = (int)( (cur_char == p[1]) ? p[2] : 0 );
583:          else
584:                      rval = (int)( p[(cur_char - 'a') + 1] );
585:
586:          if( Debug )
587:                      printf("Next state = %d\n", rval );
588:
589:          return( rval );
590: #endif
591: }
592:
593: /*--------------------------------------------------------------
594:  * State machine tables:
595:  *
596:  * There are two types of states. If the first entry is 0 then the state
597:  * is an array holding the next state, indexed by the current input
598:  * character. If the first entry is 1 then all transitions but one are to
599:  * state 0, the next character is the one exception and the transition
600:  * state follows.
601:  *
602:  * --------------------------------------------------------------
603:  */
604:
605: S(s0 )={ 0, 84,84,84,1,1,84,84,84,84,84,84,84,84,84,84,84,1,84,
606:                                       84,84,84,84,84,84 };
607:
608: S(s1 )={ 0, 0,0,2,0,7,0,21,0,0,0,23,0,30,0,0,0,33,38,45,0,0,0,0,49,0};
609: S(s2 )={ 1, 'i', 3   };
610: S(s3 )={ 1, 'p', 4   };
611: S(s4 )={ 1, 'o', 5   };
612: S(s5 )={ 1, 'c', 6   };
613: S(s6 )={ 1, 's', 87  };
614: S(s7 )={ 0, 0,0,0,0,0,0,0,0,0,0,0,8,0,0,0,4,0,10,0,16,0,15,0,0,0,19 };
615: S(s8 )={ 1, 'b', 9   };
616: S(s9 )={ 1, 'a', 55  };
617: S(s10)={ 0, 0,0,0,0,11,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,14,0,0,0,0,0 };
618: S(s11)={ 1, 'h', 12  };
619: S(s12)={ 1, 'p', 6   };
620: S(s13)={ 0, 0,0,0,0,0,0,0,0,0,0,0,0,0,81,0,0,0,81,0,0,0,0,0,0,0,0 };
621: S(s14)={ 1, 't', 87  };
622: S(s15)={ 1, 'i', 14  };
623: S(s16)={ 1, 'a', 17  };
624: S(s17)={ 0, 0,0,18,0,0,0,0,0,0,0,0,18,0,0,0,0,0,0,0,0,0,0,0,0,0,0 };
625: S(s18)={ 0, 88,0,0,0,88,0,0,0,0,0,0,88,0,0,0,0,0,88,0,0,0,0,88,0,0,88,0};
626: S(s19)={ 1, 'i', 20  };
627: S(s20)={ 1, 'l', 80  };
628: S(s21)={ 1, 'n', 22  };
629: S(s22)={ 1, 'i', 86  };
630: S(s23)={ 0, 25,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,24,0,0,0,0 };
```

C Chest

Listing Two

```
631: S(s24)={ 1, 'f', 83 };
632: S(s25)={ 0, 0,0,87,0,0,0,0,0,26,0,0,0,0,27,0,0,0,0,0,87,0,0,0,0,0,0 };
633: S(s26)={ 0, 0,0,87,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,87,0,0,0,0,0,0 };
634: S(s27)={ 1, 'o', 28 };
635: S(s28)={ 1, 'i', 29 };
636: S(s29)={ 1, 't', 89 };
637: S(s30)={ 1, 'o', 31 };
638: S(s31)={ 1, 'i', 32 };
639: S(s32)={ 0, 0,0,6,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,87,0,0,0,0,0,0 };
640: S(s33)={ 1, 'e', 34 };
641: S(s34)={ 1, 'h', 35 };
642: S(s35)={ 1, 'p', 36 };
643: S(s36)={ 1, 'a', 37 };
644: S(s37)={ 1, 'r', 87 };
645: S(s38)={ 0, 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,39,0,41,0,0,0,0,0 };
646: S(s39)={ 1, 'e', 40 };
647: S(s40)={ 0, 0,0,0,0,0,0,0,0,0,0,83,0,83,0,0,0,0,0,0,0,0,0,0,0,0 };
648: S(s41)={ 1, 'o', 42 };
649: S(s42)={ 1, 'i', 43 };
650: S(s43)={ 1, 'c', 44 };
651: S(s44)={ 0, 88,88,88,88,88,88,88,88,88,88,88,88,88,88,88,88,88,88,87,88,
652:                                                88,88,88,88,88,88 };
653: S(s45)={ 1, 'n', 46 };
654: S(s46)={ 1, 'e', 47 };
655: S(s47)={ 0, 0,0,0,87,0,0,0,48,0,0,0,83,0,0,0,0,0,0,0,0,0,0,0,0,0 };
656: S(s48)={ 1, 'c', 87 };
657: S(s49)={ 0, 0,0,0,0,0,0,50,35,0,0,0,83,0,0,0,0,0,51,0,0,0,0,0,0,0 };
658: S(s50)={ 1, 'o', 87 };
659: S(s51)={ 1, 'a', 52 };
660: S(s52)={ 1, 'n', 53 };
661: S(s53)={ 0, 88,88,88,88,88,81,88,88,88,88,88,88,88,88,54,88,88,88,88,88,
662:                                                88,88,88,88,88,88 };
663: S(s54)={ 1, 'i', 82 };
664: S(s55)={ 0, 0,0,0,0,80,0,0,80,80,0,80,80,0,0,80,0,0,0,0,13,80,80,80,
665:                                                80,0);
666:
667: static char    *Suffixes[] =
668: {
669:         s0,  s1,  s2,  s3,  s4,  s5,  s6,  s7,  s8,  s9,
670:         s10, s11, s12, s13, s14, s15, s16, s17, s18, s19,
671:         s20, s21, s22, s23, s24, s25, s26, s27, s28, s29,
672:         s30, s31, s32, s33, s34, s35, s36, s37, s38, s39,
673:         s40, s41, s42, s43, s44, s45, s46, s47, s48, s49,
674:         s50, s51, s52, s53, s54, s55
675: };
676:
677: /*------------------------------------------------------------------------*/
678:
679: S(p0 )= {1, 0, 0 };
680: S(p1 )= {0, 0,2,4,6,9,0,0,13,22,0,0,26,29,39,41,45,50,0,53,57,64,
681:                                           0,0,0,0,0 };
682: S(p2 )= {1, 'e', 3 };
683: S(p3 )= {0, 0,0,81,0,0,0,0,81,0,0,0,0,0,0,0,0,0,81,0,0,0,81,0,0,0 };
684: S(p4 )= {1, 'o', 5 };
685: S(p5 )= {0, 0,0,0,0,0,0,0,0,0,0,0,0,87,87,0,0,0,0,0,0,0,0,0,0,0,0 };
686: S(p6 )= {1, 'i', 7 };
687: S(p7 )= {1, 's', 8 };
688: S(p8 )= {0, 85,85,85,85,85,85,85,85, 0,85,85,85,85,85,85,85,85,85,85,85,
689:                                                85,85,85,85, 0,85 };
690: S(p9 )= {0, 0,0,0,0,0,0,0,0,0,0,0,0,0,0,10,0,0,0,0,0,0,87,0,0 };
691: S(p10)= {1, 'u', 11 };
692: S(p11)= {1, 'i', 12 };
693: S(p12)= {0, 81,81,81,81,81,81,81,81,81,81,81,81,81,81,81,81,81,81,81,81,
694:                                           81, 0,81,81,81,81 };
695: S(p13)= {0, 14,0,0,0,0,0,0,0,0,0,0,0,16,0,0,0,0,0,0,0,0,19,0 };
696: S(p14)= {1, 'n', 15 };
697: S(p15)= {1, 'd', 87 };
698: S(p16)= {1, 'r', 17 };
699: S(p17)= {1, 's', 18 };
700: S(p18)= {1, 'e', 87 };
701: S(p19)= {1, 'p', 20 };
702: S(p20)= {1, 'e', 21 };
703: S(p21)= {1, 'r', 84 };
704: S(p22)= {0, 0,0,0,0,0,0,0,0,0,0,0,0,86,69,0,0,0,0,0,0,0,0,0,0,0,0 };
705: S(p23)= {0, 81,0,0,0,0,81,81,0,0,0,0,81,81,0,0,0,0,0,0,0,0,0,0,0,0 };
706: S(p24)= {0, 0,0,0,0,21,0,0,0,0,0,0,0,0,0,0,0,25,0,0,0,0,0,0,0,0 };
707: S(p25)= {1, 'o', 87 };
708: S(p26)= {1, 'e', 27 };
709: S(p27)= {1, 'x', 35 };
710: S(p28)= {1, 'i', 87 };
711: S(p29)= {0, 30,0,0,0,0,0,0,0,36,0,0,0,0,0,0,0,0,0,0,37,0,0,0,0,0 };
712: S(p30)= {0, 0,0,31,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,33,0,0,0,35,0,0 };
713: S(p31)= {1, 'r', 32 };
714: S(p32)= {1, 'o', 83 };
715: S(p33)= {1, 'h', 34 };
716: S(p34)= {1, 'e', 82 };
717: S(p35)= {1, 'i', 82 };
718: S(p36)= {1, 'n', 35 };
719: S(p37)= {1, 'l', 38 };
720: S(p38)= {1, 't', 65 };
721: S(p39)= {1, 'o', 40 };
722: S(p40)= {1, 'n', 86 };
```

C Chest (Listing continued, text begins on page 20)

Listing Two

```
723: S(p41)= {0, 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,42,43,0,0,0,0 };
724: S(p42)= {1, 't', 87 };
725: S(p43)= {1, 'e', 44 };
726: S(p44)= {1, 'r', 86 };
727: S(p45)= {1, 's', 46 };
728: S(p46)= {1, 'e', 47 };
729: S(p47)= {1, 'u', 48 };
730: S(p48)= {1, 'd', 49 };
731: S(p49)= {1, 'o', 83 };
732: S(p50)= {1, 'u', 51 };
733: S(p51)= {1, 'a', 52 };
734: S(p52)= {1, 'd', 87 };
735: S(p53)= {0, 0,0,0,0,54,0,0,0,0,0,0,0,0,0,55,0,0,0,0,0,56,0,0,0,0,0 };
736: S(p54)= {1, 'm', 28 };
737: S(p55)= {1, 'm', 18 };
738: S(p56)= {0, 0,0,87,0,0,0,0,0,0,0,0,0,0,0,0,0,20,0,0,0,0,0,0,0,0,0,0 };
739: S(p57)= {0, 0,0,0,0,0,0,58,0,0,0,0,0,0,0,0,0,0,61,0,0,0,0,0,0,0,0,0 };
740: S(p58)= {1, 'e', 59 };
741: S(p59)= {1, 'r', 60 };
742: S(p60)= {1, 'e', 87 };
743: S(p61)= {0, 62,0,0,0,0,0,0,0,68,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0 };
744: S(p62)= {1, 'n', 63 };
745: S(p63)= {1, 's', 23 };
746: S(p64)= {1, 'n', 66 };
747: S(p65)= {1, 'i', 83 };
748: S(p66)= {0, 85,85,85,70,85,85,85,85, 0,85,85,85,85,85,85,85,85,
749:                                  85,85,85,85,85,85,85,85 };
750: S(p67)= {1, 'r', 87 };
751: S(p68)= {0, 81,0,0,0,0,81,0,0,0,0,0,0,0,0,0,0,0,0,0,81,0,0,0,0,0 };
752: S(p69)= {0, 85,85,85,85,85,85,85,85,85,85,85,85,85,85,85,85,85,
753:                                  85,24,85,85,85,85,85,85 };
754: S(p70)= {1, 'e', 67 };
755:
756:
757: static char    *Prefixes[] =
758: {
759:         p0,   p1,   p2,   p3,   p4,   p5,   p6,   p7,   p8,   p9,
760:         p10,  p11,  p12,  p13,  p14,  p15,  p16,  p17,  p18,  p19,
761:         p20,  p21,  p22,  p23,  p24,  p25,  p26,  p27,  p28,  p29,
762:         p30,  p31,  p32,  p33,  p34,  p35,  p36,  p37,  p38,  p39,
763:         p40,  p41,  p42,  p43,  p44,  p45,  p46,  p47,  p48,  p49,
764:         p50,  p51,  p52,  p53,  p54,  p55,  p56,  p57,  p58,  p59,
765:         p60,  p61,  p62,  p63,  p64,  p65,  p66,  p67,  p68,  p69,
766:         p70
767: };
768:
769: /*================================================================
770:  *
771:  *        MSC DEBUGGING ROUTINES (INCLUDING A MAIN MODULE):
772:  */
773:
774: #ifdef DEBUG
775: bprint( b , e )
776: ATYPE    *b, *e;
777: {
778:         /*     Print all characters between b and e inclusive    */
779:
780:         while( b <= e && *b )
781:                 putchar( *b++ & 0x7f );
782: }
783:
784: /*----------------------------------------------------------------*/
785:
786: char    *sgets( prompt, buf, maxch )
787: char    *prompt, *buf;
788: {
789:         /* Print prompt to stdout, then fill buf fm stdin. Get at most
790:          * maxch characters. Return pointer to string terminator (ie. the
791:          * null at the end of the string) or 0 on end of file.
792:          */
793:
794:         register int     c;
795:         register char    *start;
796:
797:         start = buf;
798:         printf( prompt );
799:
800:         while( (c = getchar()) != EOF && c != '\n' && --maxch > 0 )
801:                 *buf++ = c;
802:
803:         *buf = 0;
804:
805:         return (c == EOF && start == buf) ? (char *)0 : buf ;
806: }
807:
808: /*----------------------------------------------------------------*/
809:
810: phyphen( p )
811: ATYPE *p;
812: {
813:         /*     Print a word with all the hyphens visible
814:          */
815:
816:         printf("<");
```

```
817:
818:            for( ; *p ; putchar(*p++ & 0x7f) )
819:                    if( HAS_HYPHEN(*p) )
820:                            putchar('-');
821:
822:            printf(">\n");
823: }
824:
825: /*----------------------------------------------------------------------*/
826:
827: main(argc, argv)
828: int     argc;
829: char    **argv;
830: {
831:            char    buf[132], *p;
832:
833:            if( argc > 1 )
834:                    Debug = 1;
835:
836:            while( p = sgets( "word: ", buf, 132 ) )
837:            {
838:                    if( !*buf )
839:                            printf("\n");
840:
841:                    else if( p-- > buf )
842:                    {
843:                            hyphen( buf, p );
844:                            putchar('\t');
845:                            phyphen( buf );
846:                    }
847:            }
848: }
849:
850: #endif
```

End Listing Two

Listing Three

```
+----------------------------------------------------------------------+
|   Diagnostic output from hyphen.c when compiled with the DEBUG #defined: |
+----------------------------------------------------------------------+

word: be/cxxxe/able      <be-cxxxe-able>
word: be/hxxxh/able      <be-hxxxh-able>
word: be/sxxxi/able      <be-sxxxi-able>
word: be/wxxxk/able      <be-wxxxk-able>
word: com/xxxl/able      <com-xxxl-able>
word: con/xxxo/able      <con-xxxo-able>
word: dis/xxxu/able      <dis-xxxu-able>
word: equi/xxv/able      <equi-xxv-able>
word: ex/xxxxw/able      <ex-xxxxw-able>
word: hand/xxxxxx/able   <hand-xxxxxx-able>
word: horse/xxxxy/able   <horse-xxxxy-able>
word: hyper/xxxnt/able   <hy-per-xxxnt-able>
word: im/xxxxxxrt/able   <im-xxxxxxrt-able>
word: in/xxxxxxion/ary   <in-xxxxxxion-ary>
word: inter/xxxxen/ary   <in-ter-xxxxen-ary>
word: intro/xxxxx/nary   <in-tro-xxxxx-nary>
word: lexi/xxxxxxx/cal   <lex-i-xxxxxxx-cal>
word: macro/xxxxa/cate   <mac-ro-xxxxa-cate>
word: mathe/xxxxe/cate   <math-e-xxxxe-cate>
word: maxi/xxxxxi/cate   <max-i-xxxxxi-cate>
word: mini/xxxxxo/cate   <min-i-xxxxxo-cate>
word: multi/xxxxu/cate   <mul-ti-xxxxu-cate>
word: non/xxxxxxy/cate   <non-xxxxxxy-cate>
word: out/xxxxxxx/cial   <out-xxxxxxx-cial>
word: over/xxxxx/cious   <over-xxxxx-cious>
word: pseudo/xx/scious   <pseu-do-xx-scious>
word: quad/xxxxx/cient   <quad-xxxxx-cient>
word: semi/xxxxx/dent    <semi-xxxxx-dent>
word: some/xxxxxxx/ful   <some-xxxxxxx-ful>
word: sub/xxxxxxxl/ize   <sub-xxxxxxxl-ize>
word: super/xxxxa/late   <su-per-xxxxa-late>
word: there/xxxxe/late   <there-xxxxe-late>
word: trans/axxxi/late   <trans-ax-xxi-late>
word: trans/fxxxo/late   <trans-fxxxo-late>
word: trans/gxxxu/late   <trans-gxxxu-late>
word: trans/lxxxy/late   <trans-lxxxy-late>
word: trans/mxxxx/less   <trans-mxxxx-less>
word: tri/axxxxxxxx/ly   <tri-ax-xxxxxxx-ly>
word: tri/fxxxxxxxx/ly   <tri-fxxxxxxx-ly>
word: tri/uxxxxxxxx/ly   <tri-ux-xxxxxxx-ly>
word: under/xxxxx/ment   <un-der-xxxxx-ment>
word: un/xxxxxxxx/ness   <un-xxxxxxxx-ness>
word: /xxxxx/ogy         <xxxxx-ogy>
word: /xx/rapher         <xx-rapher>
word: /xxx/raphy         <xxx-raphy>
word: /xxx/scope         <xxx-scope>
word: /xx/scopic         <xx-scopic>
word: /xxx/scion         <xxx-scion>
word: /xx/sphere         <xx-sphere>
word: /xxxxx/tal         <xxxxx-tal>
word: /xxxx/tial         <xxxx-tial>
word: /xxxx/tion         <xxxx-tion>
word: /xx/tional         <xx-tion-al>
word: /xxxx/tive         <xxxx-tive>
word: /xxxx/ture         <xxxx-ture>
word: /xxxxxl/ing        <xxxxxl-ing>
word: /xxxxl/ing         <xxxxl-ing>
word: /xxxg/gling        <xxxg-gling>
word: /xxxa/bling        <xxxa-bling>
```

(Continued on next page)

C Chest   (Listing continued, text begins on page 20)
Listing Three

```
word: /xxxck/ling          <xxxck-ling>
word: /xxxg/kling          <xxxg-kling>
word: /duck/ing            <duck-ing>
word: /fit/ting            <fit-ting>
word: /hiss/ing            <hiss-ing>
word: /swell/ing           <swell-ing>
word: /fic/tion            <fic-tion>
word: /fic/tional          <fic-tion-al>
word: /fic/tionalize       <fic-tion-al-ize>
word: /fic/tionalized      <fic-tion-al-ized>
word: /fic/tionalizes      <fic-tion-al-izes>
word: /uuuccuuu/           <uuuc-cuuu>
word: /uuuquuuu/           <uuu-quuuu>
word: /uuuckuuu/           <uuuck-uuu>
word: /uuuccuuu/           <uuuc-cuuu>
word: /uuuchchuu/          <uuuch-chuu>
word: /uuughghuu/          <uuugh-ghuu>
word: /uuuphphuu/          <uuuph-phuu>
word: /uuushshuu/          <uuush-shuu>
word: /uuuththuu/          <uuuth-thuu>
word: /uuuxxxxxx/          <uuux-xxxxx>
word: /uuullaxxxx/         <uuul-lax-xxx>
word: /uuussexxxx/         <uuus-sex-xxx>
word: /uuufteee/e          <uuuf-teeee>
word: /uuuxxer/            <uuux-xer>
word: /uuuxxer/s           <uuux-xers>
word: /uuuxxag/e           <uuux-xage>
word: /uuuxxage/s          <uuux-xages>
word: /uuuxxest/           <uuux-xest>
word:
word:    <Test repetetive suffix and prefix removal:>
word:
word: disiminnonoverun/en/aryfulizelesslymentableness
              <dis-im-in-non-over-un-en-ary-ful-ize-less-ly-ment-able-ness>
word:
word: super/califragilisticexpialido/cious
              <su-per-califragilis-ticex-pialido-cious>
word:
word:    <All words following this point test exceptions to>
word:    <the hyphenation rules and shouldn't be hyphenated.>
word:
word: /uuussx/             <uuussx>
word: /uuullx/             <uuullx>
word: /uuuller/            <uuuller>
word: /uuuller/s           <uuullers>
word: /uuubluu/            <uuubluu>
word: /uuubruu/            <uuubruu>
word: /uuucluu/            <uuucluu>
word: /uuucruu/            <uuucruu>
word: /uuuchluu/           <uuuchluu>
word: /uuuchruu/           <uuuchruu>
word: /uuudguu/            <uuudguu>
word: /uuudruu/            <uuudruu>
word: /uuufluu/            <uuufluu>
word: /uuufruu/            <uuufruu>
word: /uuughtuu/           <uuughtuu>
word: /uuugluu/            <uuugluu>
word: /uuugruu/            <uuugruu>
word: /uuuknuu/            <uuuknuu>
word: /uuulkuu/            <uuulkuu>
word: /uuulquu/            <uuulquu>
word: /uuunchuu/           <uuunchuu>
word: /uuunkuu/            <uuunkuu>
word: /uuunxuu/            <uuunxuu>
word: /uuuphruu/           <uuuphruu>
word: /uuupluu/            <uuupluu>
word: /uuupruu/            <uuupruu>
word: /uuurkuu/            <uuurkuu>
word: /uuuspuu/            <uuuspuu>
word: /uuusquu/            <uuusquu>
word: /uuutchuu/           <uuutchuu>
word: /uuutruu/            <uuutruu>
word: /uuuthruu/           <uuuthruu>
word: /uuuwhuu/            <uuuwhuu>
word: /uuuwluu/            <uuuwluu>
word: /uuuwnuu/            <uuuwnuu>
word: /uuuwruu/            <uuuwruu>
word: /uuufter/            <uuufter>
word: /uuufter/s           <uuufters>
word: /uuuftag/e           <uuuftage>
word: /uuuftage/s          <uuuftages>
word: /uuuftest/           <uuuftest>
word: /uuulder/            <uuulder>
word: /uuumper/            <uuumper>
word: /uuunder/            <uuunder>
word: /uuunger/            <uuunger>
word: /uuunser/            <uuunser>
word: /uuunter/            <uuunter>
word: /uuurger/            <uuurger>
word: /uuurner/            <uuurner>
word: /uuurter/            <uuurter>
word: /uuuster/            <uuuster>
```

End Listings

Circle **no. 77** on reader service card.

41

A Threaded-Code Microprocessor Bursts Forth

Subroutines Without Performance Anxiety

by Leo Brodie

Subroutine calls in one clock cycle are one advantage as Forth goes to silicon.

The Novix NC4000 processor has arrived, and it threatens to shake most of our beliefs about microprocessor design. What implications does this new technology have for the art of programming? This article will explain the techniques used by Forth and by the NC4000 microprocessor, and explore the impact of this technology on software design.

Subroutines: The More the Better

Most programmers recognize the importance of dividing a large, complex program into a group of smaller, manageable pieces. The use of smaller "modules" is advantageous for the following reasons:

• smaller pieces of code are easier to think about, and therefore easier to design and to write;
• smaller pieces of code are easier to debug; the likelihood of introducing a bug is reduced, and, if one does appear, it will be easier to find;
• different parts of a program may be assigned to different programmers;
• subroutines can be used to hide information about things that might change in subsequent revisions; if a change must be made to a program, only the subroutine has to be changed.

Unfortunately, most conventional high-level languages (HLLs) are not optimized for the invocation of subroutines. One problem is the loss of time caused by saving registers before a jump to a subroutine and restoring them afterwards. This has been estimated as 40% of program execution time. Even more time-consuming is the invisible but significant code needed to pass arguments to and from the subroutine. At the machine level as well, processors are not generally optimized for invoking subroutine calls.

For these reasons, although the use of many, small subroutines is an ideal, nevertheless, in practice, programmers tend to writer fewer, larger subroutines. When efficiency is at stake, it's hard to appreciate the elegance of subroutines.

There is a high-level programming language optimized for subroutines: Forth. And now, a single-chip microprocessor is available that applies Forth's optimization techniques directly to a silicon architecture. The result is a machine that can execute high-level subroutine calls and returns in a single clock cycle!

Threaded Code

Let's begin by examining what the inherent architecture of Forth is, regardless of which processor it runs on. Forth is implemented using a technique called "threaded code." Suppose we define a word in Forth called PROCESS, consisting of three operations.

```
:   PROCESS        STEP1  STEP2  STEP3  ;
```

In this definition, the name of the word (subroutine) is PROCESS. It consists of three lower-level subroutines. The

colon begins the definition, while the semicolon ends it.

Figure 1 (page 44) shows what this definition looks like when compiled into memory on an emulated Forth system. Such a definition consists of the following parts:

• The name field contains an ASCII representation of the word's name, so that the word can subsequently be found and executed.
• The link field points to the previous definition in the linked list of dictionary entries. (A search through the dictionary for a given word begins with the most recently defined word, and, using the link field for each word, proceeds backwards through the chain or chains of words.)
• The code pointer field points to machine-executable code that is generic to each "family" of definitions. As yet we have only discussed colon definitions; we will introduce other types later.
• The remainder of the definition is a list of addresses of other routines. Depending on the processor, these addresses are usually 16- or 32-bit numbers. In this example, the addresses of the routines STEP1, STEP2, and STEP3 are listed. At the end of the list is the address of another routine called EXIT, which is compiled by the semicolon and which terminates the definition when it is executed.

The unusual thing about threaded code is the absence of any CALL instructions—*everything* is a subroutine. Definitions consist simply of addresses of other definitions. As we'll see, even data structures are defined as dictionary entries, so that their addresses appear in definitions when referenced.

If you expand your imagination a bit, you can view these addresses as actual *instructions*, since each address is unique. This is exactly the way they are viewed by the Forth machine, and in particular, by the routine called the "address interpreter," also known as NEXT.

It is the function of the address interpreter to execute each "Forth instruction" in turn. This function is analogous to the operation of a machine processor wherein each opcode is interpreted and the program counter advanced. An "interpreter pointer," called I (usually a register in the real processor) operates analogously to the processor's program counter: during execution of any Forth instruction, I points to the next instruction to be executed; it is incremented by the address interpreter.

To begin with the simplest case possible, let's assume that each of the definitions STEP1, STEP2, and STEP3 is defined in machine code—they are not high level Forth definitions. (Forth allows routines to be coded in high-level code or in machine code without distinction; in practice, however, only a handful of time-critical or hardware-dependent routines must be machine-coded.) Here's what happens, in somewhat simplified form for now. I is pointing to the first address in the definition of PROCESS; this is the address of STEP1. The address interpreter, NEXT, first advances I to the next cell (the one containing the address of STEP2), then jumps to STEP1. STEP1 is a machine code definition, executed in the normal way by the

Figure 1
A compiled definition in the dictionary of an emulated Forth system

Figure 2
The code pointer of an assembler definition (e.g., STEP1) points to its own definition. The code pointer of a high-level definition (e.g., PROCESS) points to a prologue routine called CALL.

Figure 3
When STEP2's CALL routine is invoked, it saves I (the interpreter pointer) on the return stack, puts the address of the body of STEP2's definition in I, then invokes NEXT.

underlying processor. The final piece of code in STEP1 is NEXT once again (either in-line or as a jump instruction, depending on whether performance is a more or less important consideration than memory usage in the particular hardware configuration). This second NEXT increments I to the third cell, then jumps to STEP2, and so on.

Each machine-coded definition (called a primitive) must end by executing the code for NEXT to keep the Forth machine running.

In the preceding discussion, we omitted one important detail. The address interpreter does not jump to the address of each new instruction, but rather to the *code pointed to by the code field* of each new instruction—one added level of indirection. Figure 2 (at left) shows this relationship.

In the case of a machine code definition, the code field points to the body of the definition itself, where the actual machine instructions have been compiled. In the case of all colon (high-level) definitions, however, the code field points to a single routine in memory that we will name CALL. Essentially this routine handles the job of nesting definitions, so that one high-level word may invoke another word, which in turn invokes another, and so on.

Here is what CALL does, step by step. Using our previous example, let's now suppose that STEP2 has been defined as a colon, or high-level, definition. We are executing PROCESS, and have just completed STEP1. I (the interpreter pointer) is pointing to the address of STEP2. The address interpreter advances the pointer in anticipation of the return, then gets the address of STEP2. Using this address, it jumps to the address pointed to by the code pointer field of STEP2. Since STEP2 is a high-level definition, the routine at this address is CALL. This routine saves I on a special stack for return addresses called the return stack. It then places into I the address of the beginning of the instruction list in STEP2 (see Figure 3, at left). This done, CALL executes NEXT. Thus the Forth machine continues as before, but has nested in a level.

The opposite of the CALL function is EXIT, which terminates each high-level definition. When the EXIT in STEP2 is encountered, it pops the top of the return stack back into the interpreter pointer, thus resuming execution of PROCESS at the address of STEP3.

Because of the extra level of indirection offered by the code pointer field, the architecture we've described has been called indirect threaded code (ITC).

The value of the code pointer field is that each word knows what kind of procedure it is, and how to invoke itself. This is what makes it possible in Forth to design a routine first in high-level code, then to rewrite it in machine code without changing any code elsewhere that invokes the routine.

Each family of data structures in Forth uses a unique prologue routine, pointed to by the code fields of all members of that family. A simple example is the constant, as illustrated by Figure 4 (page 45). A Forth constant is a dictionary entry, possessing the same type of header structure as all dictionary entries. Its body consists of a code pointer for all constants, along with the value of the

44

constant. The code pointer points to a routine, which we will call "do-constant," whose job it is to fetch the value from the body of the constant, then invoke NEXT.

The Forth Engine in Silicon

In indirect threaded code implementations of Forth as just described, the invoking of words is elegant and more efficient than the calling of subroutines in other high-level languages. However, because conventional microprocessors are not optimized for ITC, running the Forth machine involves some performance overhead. This overhead consists of the time spent executing NEXT after each machine instruction, plus the time spent for each call and return. Running high-level Forth versus straight machine code can increase execution time by as much as 100 percent (which is still better than other high-level languages).

A primary goal of the NC4000 architecture is to allow the execution of threaded code with as little overhead as possible. This goal is achieved first by eliminating NEXT as a software routine. Instead, NEXT has become the *background* to the chip's operation. The overhead of NEXT is reduced from many clock cycles to less than one (see Figure 5, at right).

Secondly, the NC4000 reduces the normal overhead of the CALL operation. In fact, it achieves this goal to the utmost degree possible: a subroutine call executes in a *single clock cycle*. Compare this with the 15 clocks required by the 8086 to effect a machine-level subroutine call. A high-level language such as C will require still more instructions to effect a jump to subroutine.

How does the NC4000 achieve these results? One of the bits (the MSB) of the 16-bit opcode is reserved exclusively to indicate a subroutine call. For all ordinary machine instructions, this bit is one; but if this bit is zero, the NC4000 effects a call to the subroutine whose address is specified by the remaining 15 bits. Thus, colon definitions appear as compiled object code as shown in Figure 6 (at right).

As you can see, compiled object code for the NC4000 consists of a mixture of actual machine instructions and single-instruction subroutine calls to definitions. In other words, the NC4000 executes high-level Forth threaded code as its native machine language.

Of actual machine instructions the NC4000P boasts over 170 combinations, not including permutations of register addressing. These include the usual Forth primitives for data-stack manipulation, arithmetic, memory, register, and I/O fetches and stores, jumps and loops—even multiply, divide, and square root steps.

The use of the high-order bit enables the chip's logic to decode each instruction early in the clock cycle. If the instruction is a CALL, then the address represented by the remaining 15 bits will appear on the address bus in time for the appropriate instruction (in the called routine) to be latched and executed on the next cycle.

The NC4000 does even better with the EXIT routine. Another of the 16 bits in the opcode (octal 40) is reserved exclusively to indicate the return operation. The return can occur simultaneously with other operations of the

Figure 4
A Forth constant is a normal dictionary entry whose code pointer points to a routine common to all constants.

Figure 5
Comparison of overhead for NEXT

Figure 6
A definition list on the NC4000.

ALU. In a single clock cycle, a word can perform its last operation and return to the word that invoked it. Thus, in most cases, there is *no overhead* for the EXIT operation.

One consequence is that program space is limited to 32K addresses. Since the NC4000 uses *word addresses*, not byte addresses, the program may occupy 64K bytes of memory. Given the extreme compactness of threaded-code systems in general, and of this instruction set in particular (each instruction can perform up to five operations simultaneously), 64K of program memory will contain a considerable amount of code compared to traditional systems. Since memory operators use full 16-bit addresses, program size may be further compacted by placing data buffers and tables above the 32K word boundary instead of in the program region. For those who still aren't satisfied, Novix is currently considering development of a 32-bit version of the NC4000 architecture.

A necessary ingredient in this design is the implementation of a return stack in hardware. The return stack can be expressly manipulated by the programmer, but it is also controlled automatically by the CALL and EXIT operations of the processor. The NC4000 features a unique parallel architecture that allows the return stack to operate simultaneously with other CPU operations.

Notice in Figure 6 that the code field is no longer needed, because Forth is the native machine code of the NC4000. The CALL routine, for instance, which serves as the prologue for colon definitions, now exists in processor logic, and is invoked whenever bit 15 is zero. In the case of constants, the pointer to the old "do-constant" routine can now be replaced by a literal-fetch instruction.

In fact, each dictionary entry, whether a definition, a constant, variable, array, or what-have-you, is most efficiently compiled as a definition and directly executed. Thus, the Novix processor uses "direct-threaded-code" (DTC), not ITC.

To summarize, the NC4000 realizes the Forth machine directly in silicon logic, without even resorting to microcode. As a result, it supports a subroutine call and return with an overhead of one clock cycle in most instances, and never more than two. Comparable CALL and RET instructions on the 8086 require a total of 34 clock cycles (and each 8086 cycle is two to five times slower than that of the Novix prototype).

The Software Implications

What does all of this mean to the programmer? It is easily demonstrated that the NC4000 runs faster than conventional micros. Because Forth instructions execute in a single clock cycle, the chip runs Forth code about 100 times faster than Forth running on a conventional processor. A benchmark using the Sieve of Eratosthenes reveals that the NC4000 runs Forth over 10 times faster than the 68000 runs its own machine code. (This benchmark was made on the NC4000P, the initial CMOS gate-array prototype; future revisions will increase the speed considerably.)

But these benchmarks only begin to show the potential power of the NC4000 architecture. A benchmark such as the Sieve represents a straightforward algorithm, which is

best coded in the linear terms of conventional processors. There is nothing to be gained by decomposing such an algorithm into smaller modules. Although a 10-times speed improvement is nothing to be ashamed of, the Sieve benchmark doesn't let the NC4000 show its true stuff.

The real wonder of the Novix chip is that it allows execution of an elegant, high-level, modular language directly in the logic of the CPU.

Forth and the NC4000 offer three distinct benefits in terms of programming methodology. The first is the interchangeability of a word without affecting any code that invokes the word. The second is the ability to use words, which tend to be shorter and simpler, instead of subroutines which tend to be longer. The third is that more efficient subroutines encourage greater freedom in the use of modular programming. Let's examine these points more closely.

Interchangeability

In the world of software development, it's no longer enough just to get a program running. The program must also be able to be maintained and easy to change. Why would a program ever need to change, once it's running? You may need to upgrade the program to run on newer equipment. You may need to add more sophisticated features to keep up with the competition. In fact, most software groups find themselves writing families of programs; that is, they write many versions of related programs in their application field, each a variant on an earlier program.

To be maintainable, a program must be easy to read, easy to understand, and, most importantly, structurally durable. By this I mean that you can rewrite a section of the program without having the whole thing collapse.

Here is an example of how Forth encourages structurally durable code. Suppose we have a variable called STATUS. In Forth this is defined as

    VARIABLE STATUS

One component of our application, a device driver, for example, writes into STATUS, another component reads STATUS. References to STATUS are present throughout the program.

In Forth, variables are locations that can be acted upon. For instance, in the phrase

    STATUS   @

the word @ (pronounced "fetch") fetches the value of STATUS. In the phrase

    STATUS   !

the word ! (pronounced "store") stores a given value into STATUS. In either phrase, the execution of the word STATUS is the same: it merely returns the address of the location of its data. This address is then used by the @ or ! operators (or by any other memory manipulation operation).

47

Now suppose we must write the next generation of the program. This time we are monitoring ten devices instead of just one and we must handle a STATUS variable for each of them. What we need is a ten-cell array and a pointer to the cell in the array for the current device. In Forth we merely define:

```
CREATE STATUSES   20 ALLOT   ( 10 cells)
VARIABLE DEVICE ( number of current device)
:   STATUS   ( – – adr)
  DEVICE @ 2*      STATUSES + ;
```

On the first line we have defined a 10-cell array (20 bytes) called STATUSES. On the second line we've created a variable that will contain the number of the current device (between 0 and 9). On the third line we've defined the word STATUS to fetch the number of the current device, multiply this by two (to compute the offset into the array in bytes), then add it to the address of the top of the array. The result is the address of the cell containing the status for the current device.

Now the other components of the application may write

    STATUS @

and

    STATUS !

as before, even though we have changed the definition of STATUS from that of a variable to a colon definition—from a data structure to a procedure. Forth allows us to hide the details of how STATUS is defined from the code that uses it. What appears to be a thing (a variable) to the original code is actually defined now as an action.

Words, Not Subroutines

Strictly speaking, Forth words are no different than ordinary subroutines. However, because they are optimized, and because they require no special calling command in the code that invokes them, Forth encourages a finer granularity in the decomposition of an application. In other words, Forth words tend to be much shorter and more numerous than typical subroutines in other languages. This characteristic imparts a style that is unique to Forth, and it is one of its most difficult characteristics to describe.

Here is an example of a collection of subroutines that might be written in a conventional language:

    ENABLE-LEFT-MOTOR
    ENABLE-RIGHT-MOTOR
    DISABLE-LEFT-MOTOR
    DISABLE-RIGHT-MOTOR
    ENABLE-LEFT-SOLENOID

ENABLE-RIGHT-SOLENOID
DISABLE-LEFT-SOLENOID
DISABLE-RIGHT-SOLENOID

In Forth one might replace these eight subroutines with these six words:

ENABLE
DISABLE
LEFT
RIGHT
MOTOR
SOLENOID

Now our syntax reduces to simple, English words, combined in useful phrases:

ENABLE LEFT MOTOR
ENABLE RIGHT MOTOR
Etc.

Not only do we reduce the number of routines from eight to six, we also make each word's definition much shorter and simpler. The following listing shows how little code we might need to define the syntax:

−1 CONSTANT ENABLE (all "one"s)
0 CONSTANT DISABLE (all zeros)

1 CONSTANT LEFT (bit mask)
2 CONSTANT RIGHT (bit mask)
: MOTOR (action side − −) 'MOTOR SET-BIT ;
: SOLENOID (action side − −) 'SOLENOID SET-BIT ;

We have defined the word SET-BIT to require three arguments: a flag, either all ones or all zeros; a bit-mask, with a "1" indicating the desired bit; and an address. SET-BIT sets the bit indicated by the mask at the given address according to the given flag. The words 'MOTOR and 'SO-LENOID provide the hardware addresses of the ports that communicate to the motor and solenoid respectively; bits zero and one of these addresses control the left and right devices, respectively.

In well-written Forth code, the average definition consists of only seven words, and fits within a single line of code (or two lines if a long comment is used). Few other languages encourage you to create subroutines called LEFT and RIGHT, merely to act as "adjectives." Nor would you find routines such as MOTOR that consist of only two other words.

In conventional programming languages, subroutines tend to be longer and more complex, covering many possible options. A conventional subroutine will contain a number of conditional branches, so that it can decide how to behave depending on the desired function. Typically, many of these decisions shouldn't have to be made at runtime; they arise only because the subroutine is too gener-

al. It has not been decomposed sufficiently.

The product of good decomposition is simplicity. Simplicity is the key not only to efficient, compact programs, but also to programs that are easy to write and, therefore, easy to test and debug. They are also easy to understand and, therefore, easy to maintain.

Modularization

The third way that optimized subroutines can enhance program maintainability is by allowing the program designer to break the application into very small, useful modules. Modularization is the key to well-written Forth code.

Modularization is more than merely chopping code up into small pieces. There is no way to automate good modularization. It is something that arises from the programmer's intimate understanding of the problem being solved. Allowing the problem to decompose itself naturally is a critical part of the early design phase—one that requires the programmer's experience, sensitivity, and artistry.

The best use of modularization is to decompose the application into components. A component is a resource, such as a UART driver, a queue, a linked list, a stack, etc. Each resource will typically consist of a number of functions and data structures.

The idea of component programming is to define all these functions and data structures in the same place, not scattered throughout the program as they are needed. The primary advantage is the ability to change the component if necessary by editing code in only one location.

In Forth, component programming means creating a set of names (words) to *represent* the functions and structures of each component. The set of words veil the internal algorithms and structures, the "how it works." They present only a conceptual model of the component described in natural language, the "what it does."

These words then become the language for describing the data structures and algorithms of components written at a higher level. The "what" of one component becomes the "how" of a higher-level component. A Forth application is nothing but a collection of components.

Thus, Forth programming consists of extending the root language toward the application, providing new commands that can be used to describe the problem at hand. Forth is an environment for creating "application-oriented languages," i.e., programming languages designed for particular applications. Examples include robotics languages, control languages, data acquisition languages, and music languages.

Conclusion

Charles Moore, the creator of Forth and chief architect of the NC4000 processor, has claimed that the Forth chip represents "a landmark in the evolution of hardware and software. This article has only tapped the surface of the many revolutionary hardware features of the NC4000. Its high-speed computation and I/O characteristics have made the processor immediately interesting to engineers concerned with bandwidth in applications such as telecommunications and signal processing.

Its virtues are not limited to hardware features and real-time applications, however. I hope that this article has at least suggested some of the elegance of the processor's native language, and that you will have a chance to experience its unique style on your own.

Note

For more information on threaded code languages, refer to: "An Architectural Trail to Threaded-Code Systems," by Peter M. Kogge, in *Computer*, March 1982; also to *Threaded Interpretive Languages* by Ronald Loeliger, McGraw-Hill/Byte books. For further reading on the Novix processor, see "Fast processor chip takes its intructions directly from Forth," *Electronic Design*, March 21, 1985, or *Programmer's Introduction to the Novix NC4000P Microprocessor* available from Novix, 10590 N. Tantau Ave. , Cupertino, CA 95014. For a discussion of Forth methodology, see *Thinking Forth* by Leo Brodie (Prentice-Hall, 1984).

DDJ

Design of a Forth Target Compiler

by Howard H. Robinson,
Philip D. Morse II and
Sidney A. Bowhill

Forth is a programming language that is both flexible and extensible. It also generates threaded code that is extremely compact and executes rapidly. Forth also allows the programmer to stay within a higher-level programming environment, while at the same time making use of all of the low-level capabilities of the computer. These are ideal characteristics for a language to be used in the development of programs for machines with their code in PROM.

We had been developing code for PROM-based applications for several years. Primarily we used assemblers. After we became acquainted with FIG-Forth, we realized that it had the potential to be an excellent programming environment for the development of code for PROM applications. There were, however, two serious drawbacks. Firstly, there are large

essary for the execution of the application program. It also directs any code that must be resident in RAM to the RAM area of the target machine.

The General Design of the CICFIG Recompiler

The program described in this article is a recompiler of the type specified above. It allows programs to be developed in Forth on a host machine and recompiled to produce the target application code. It was developed by the Central Illinois Chapter of the Forth Interest Group (CICFIG) and works in conjunction with the native host Forth compiler.

To produce the machine executable target code, source code for the application program is compiled by the standard Forth compiler within the host system. The recompiler (itself already compiled into the host

By recompiling Forth programs, you can develop tight, PROMable code that executes at the top speed the target microprocessor will allow.

sections of the compiled threaded Forth code that are not needed for the functioning of the PROM-based application, but are difficult to exclude from the code (such as the Forth compiler itself). Secondly, sections of the code must be in RAM (such as self-modifying code, variables and pseudo-registers).

What was needed to rectify these problems was a post-compiler. A post-compiler, or recompiler, selects from the compiled threaded Forth code just those segments that are nec-

H. Robinson et al., University of Illinois, 1406 W. Green St., Urbana, IL 61801

Forth system) selects parts of the host system along with the compiled application code to be included in the target code. The recompiler relocates all the addresses within this code to conform to the memory map of the target machine, supplied by the application programmer. The resulting minimum length code can be executed from PROM, and potentially can perform tasks at the maximum speed of the microprocessor. The code also conforms to the memory map of the target machine, and can stand alone (i.e. does not necessarily require an operating system or monitor).

One of Forth's greatest advantages is that its interpreter can execute new-

ly compiled code. This property greatly simplifies debugging: compiled parts of the program can be tested and debugged individually with the interpreter. To extend this advantage to the process of debugging the target application, we have written the recompiler so that the target machine code can optionally include the interpreter. When the program is fully operational, the minimum code can be generated by excluding the interpreter (and the verb headers required for the interpreter to operate). In addition, once the interpreter has been excluded, the user has access only to those parts of the program allowed by the application programmer; the fact that the program is compiled Forth code should be completely transparent to the user.

Compiled Forth Code

The recompiler has as input the application program already compiled by the host Forth system. To understand what the recompiler must do, it is necessary to understand the structure of compiled FIG-Forth-model code. The model includes three parts: the start-up code, an inner interpreter, and the dictionary. The start-up code and the inner interpreter are contiguous in low memory in the host Forth system and thus can be treated as one block that must be included in the target machine code.

The start-up code initializes the registers required for the operation of the Forth inner interpreter and then jumps to the inner interpreter. The inner interpreter uses these registers to arrange the sequential execution of verbs from the dictionary.

The dictionary is a set of verbs in one linked list. Each verb defines a sequence of operations to be executed by the machine. Each verb has three parts: a header, a code field, and a parameter field. The header specifies the name of the verb (so the verb can be identified within the dictionary), a link address (so the dictionary forms one linked list of verbs), and certain special bits that are signals to the compiler of the host Forth system.

The code field contains the address of the machine executable code to which the inner interpreter will cause

the processor to jump when that verb is executed at run-time. The parameter field contains the information that is necessary for the function of the verb to be accomplished. This may be data, as in the case of variables (and arrays and constants), or a list of addresses of other verbs that are to be performed sequentially (e.g. a colon definition). Alternatively, the parameter field may contain machine executable code that performs the function of the verb directly. In this case, the code field contains the address of the parameter field.

Since Forth code consists largely of lists of addresses, we have devised a novel algorithm for identifying all addresses in the host Forth system that may need to be redirected for inclusion in the target machine code. The procedure calls for two complete copies of the host Forth system, including the compiled application program. These two Forth systems are identical except that one Forth system has its origin exactly 201 Hex locations higher in memory than the other (i.e. one has its origin at 100H and the other at 301H). When equivalent locations in both images are compared, all numbers (pairs of bytes) differing by 201H must be addresses. This algorithm identifies all absolute addresses, whether they are embedded in machine executable code, or used as data (such as the list of code field addresses in the parameter field of a colon definition).

Forth Systems for Target Recompilation

The recompiler described here is a program written entirely in high level Forth (MVP-Forth). The recompiler is itself compiled into an existing Forth system. The application program is also compiled into the same host Forth system. The recompiler examines the compiled application code and determines all code segments necessary for the proper functioning of the application program in the target machine. A file is generated that contains a copy of those necessary code segments. All addresses in the code segments in that file have been corrected to reflect the squeezing and the relocation of the host

Forth system to generate the target code image. This file then contains the code for the application program, which can be loaded into the target machine and executed.

The code generated by the recompiler is intended to be executed on a target machine with a microprocessor that is code-compatible with the host microprocessor; thus, the recompiler does not provide the function of a cross-compiler. At present, the recompiler can only generate code to be executed on either an 8080/Z80 compatible or an 8086/88 microprocessor based target machine. Development machines for these two families of microprocessors are common and can be quite modest (for instance, only one floppy disk drive is required).

We have considered adapting other Forths to allow the use of other microprocessors (such as the 6502). Time limitations, however, have precluded this. In addition, not all Forth systems are easily adapted for recompilation of this kind. The code to be included by the recompiler in the target machine code must comply with the following three rules.

Rule #1: All non-addresses in the two compiled Forth images discussed above must be equal. Thus, data storage areas, such as arrays, must be initialized. This rule is only relaxed in the case of variables, due to the possible use of certain system variables in both the host Forth system and in the target machine code (such as PREV, USE and SEC).

Rule #2: All inter-verb relative addressing mode jumps are forbidden (intra-verb relative jumps are allowed). This rule is necessary because the algorithm for detecting addresses does not detect relative jump addresses. The versions of the target compiler available from CICFIG (see the section on Availability, page 68) have all the inter-verb relative jumps in their kernels removed.

Rule #3: Both the host and the application compiled code must not contain machine code jumps and calls from one code level verb to another code level verb. This practice, common in FIG-Forth model systems, can be eliminated by recoding certain parts of the kernel.

MVP-Forth implemented on CP/M-80 and MSDOS machines seemed to conform most nearly to all our requirements. In order to make it comply with the above stated rules, we changed the kernel of the MSDOS version of MVP-Forth to replace all relative jumps to the inner interpreter with absolute jumps. The verbs 0BRANCH, </LOOP> and <LOOP> were recoded to remove jumps to BRANCH. A relative jump in <+LOOP> to <LOOP> was changed to an absolute jump. Both <VOCABULARY79> and <VOCABULARYFIG> were changed to contain an absolute jump to DODOES. DOES> was changed to put in an absolute jump to DODOES.

The kernel of the CP/M-80 version of MVP-Forth was changed to complete the following verbs without inter-verb jumps: −, 0BRANCH, <, </LOOP>, <?TERMINAL>, <CR>, <KEY>, <LOOP>, SECREAD, and U*. Two instructions

were added to the beginning of the start-up code that read and load an initial value into the cell that is to function as the pseudo-register for the return-stack pointer. This was necessary to allow this cell to be initialized in RAM when the start-up code segment is in PROM.

In both versions, the verb <+LOOP> has an inter-verb jump to BRANCH. This case is handled as an exception by the recompiler. We intended that the compiler of Forth should not be included in the resultant target machine code. Attempts to use the compiler in the target machine will have unpredictable results.

User Interface
Generation of the target machine code with this recompiler is a two-step process. The two steps are performed by the two host Forth systems TARG1 and TARG2. These two Forth systems have the recompiler already compiled into them and are identical except that they have their origins 201H locations apart.

The application programmer specifies the memory map of the target machine by loading the constants indicated in Figure 1 (page 57). These constants indicate where the origin of the target machine code is and to what value pointers to any desired RAM areas of the target machine are initialized. For simplicity, this memory map conforms to the memory map of MVP-Forth. The application programmer may rearrange this map to suit the nature of the particular problem.

The application programmer must specify the verb in the host dictionary that is to serve as the cold start verb of the target machine code. This is done by loading the constant TCOLDCFA with the code field address (cfa) of the target machine cold start verb. This verb should perform any initialization of the Forth system that may be required by the application program (for example, the "user area" in RAM may need to be initialized). The examples given in the following sections explain other initializations that may be required, depending upon the nature of the application program. The cold start verb must invoke the application program. In this way the recompiler can detect and include in the target code all verbs necessary for the operation of the cold start verb and the application program.

Upon start-up of the target machine code, the start-up routine will first be performed. This will start the inner interpreter, which executes the cold start verb and then the application program.

The procedure for the generation of target code is outlined on page 76. The recompiler is initiated with a zero on the parameter stack. This signals to the recompiler that headers should be removed from all verbs to be included in the target machine code. Alternatively, if the parameter field address (pfa) of the root verb of the application is on top of the stack, then headers will be included in the target machine code. This allows verbs in the target machine code to be "found" (by -FIND). In this case the cold start verb usually invokes an interpreter such as the verb QUIT (see example #3).

In example #1 (see screens

200–201, page 76) the recompiler is used to generate headerless code with the Forth outer interpreter excluded for a disk-based task. This program accepts HEX numbers from the keyboard and sends them to the MSDOS device PRN (usually the printer). This is useful for setting up dot-matrix printers in various modes.

Screen #200 loads a 1 into the constant MAPFILE to signal the generation of a DOS file, <codefile>.MAP, which will contain (in ASCII) the addresses and names of all verbs in the target machine code. This screen also loads the constants that define the memory map of the target machine code (see Figure 1). Loading these constants is required prior to the initiation of the recompiler.

Screen #201 contains the application program for example #1. The verb TINPUT is for numerical input from the keyboard. This avoids the Forth verb ABORT" that threads to the outer interpreter (which we are trying to exclude). LP sends a byte to the PRN device and APP is the router and root verb of the application program. Screen #201 also has the verb TCOLD which performs the initialization required to load the "user area" of RAM. The constant TCOLDCFA has been loaded with the code field address of the verb to be executed upon start-up. The code generated by this program is less than 2K.

Example #2 (screens 202–203, see page 76) demonstrates the minimum code length that can be generated by the MSDOS MVP-Forth version of the recompiler. Screen #202 loads the memory map for the target machine. Screen #203 has the verb TCOLD, which simply performs a return to the operating system. The code generated is 412 bytes, and includes the Forth start-up code, the inner interpreter, a group of mandatory verbs (which includes the run-time routine DOCOL, for instance), and INTCALL. Eleven verbs are included in the target machine code in this example.

Example #3 (screens 204–207, see page 76) demonstrates that the Forth interpreter can be included in the target machine code. The verbs in the target code must have headers so that they can be found by the outer inter-

FIGURE 1
Constants Loaded by Application Programmer to Specify Memory Map

FIGURE 2
nfa and Target Offsets List

preter. Care must be taken to ensure that the Forth virtual memory block disk verbs are not included. Screens #205–206 have the Forth outer interpreter tailored to exclude references to the disk block verbs. Screen #207 has a test application program (APP) and the cold start verb TCOLD. The vocabulary link in Forth must also be initialized. The recompiler is started by typing ' APP TARGET, rather than 0 TARGET, as in a headerless case. The cfa of the root verb of the application program is on the parameter stack; knowledge of it is necessary to the recompiler since the verb TCOLD pointed to by the constant TCOLDCFA does not thread to APP.

Example #4 (screens 208–212, page 78) shows that the verbs of the Forth virtual memory disk block handlers can be included without the outer interpreter (disconnecting certain groups of routines within the host Forth system can be a bit tricky). This program reads in a number of Forth blocks (in ASCII) and writes their contents to a DOS file (in AS-CII). This is done so that the file can be manipulated by your favorite word processor (such as WordStar; you must set the page length to 27 to manipulate 16 line Forth screens in document mode from the DOS file generated from this program).

In screen #208, the target memory map is loaded to indicate that 2 disk block buffers are to be allotted. Screens #209–210 set up the DOS file and transfer the information from the blocks to the DOS file. Trailing spaces are calculated by CHARS and removed by BLOCK>FOUT. Screen #211 has the verb TCOLD, which initializes the program. Several constants are loaded in TCOLD. It is important to be sure that the verbs referenced indirectly in TCOLD as literals (FIRST, LIMIT, #BUFF, and C/L) are included in the target machine code by direct references elsewhere in the program.

The Forth verb <R/W> has ABORT" in its definition. Consequently, ABORT, QUIT, etc. would all be included, although they would be quite useless, since the code generated is headerless. Thus, <R/W> should be revectored to a new verb which does not have this problem. Screen #212 has the verb <TR/W>, which is identical to the normal <R/W> except that the ABORT" construction has been removed. <TR/W> is re-vectored through 'R/W so that <TR/W> will be included by the recompiler rather than <R/W> (and the entire Forth outer interpreter).

The program in example #5 (screens 212–217, page 78) provides the inverse function of the program in example #4. In the DOS-ASCII files the <cr> <lf> pair is present at the end of each line of text, and unnecessary spaces are not present. These programs allow interchange between Forth blocks and DOS files. Example #5 removes the <cr> <lf> pairs and inserts spaces at the end of each line to fill-out each line of the Forth block. In screen #215, the verb TVARIABLE demonstrates how the RAM variable storage area can be used.

It may be necessary to generate code for a low-memory jump table to handle a start-up jump and interrupt jumps. In simple cases, this can be handled by hand patching. In more complex cases the recompiler can be instructed to generate a file (by setting the constant MAPFILE, 1 ' MAP-FILE !) containing the memory map (in ASCII) of all verbs in the target machine code. The code for the low-memory jump table can be generated by a program that queries the memory map file for addresses.

Error Messages and Warnings

During the process of target compilation, if the recompiler finds numbers that are not equivalent in the two Forth images and are not identifiable as addresses (or the two bytes of the data area of a variable), then an error is indicated ("Images not matched at <address>"). A brief dump is then produced, the dictionary is "unsmudged", and control is returned to the interpreter. This error is generally the result of an uninitialized data area within a verb. The application programmer must then inspect the

Circle **no. 125** on reader service card.

dump to determine the offending
verb and recode or initialize the verb.
After repairing the problem, the tar-
get compilation process must be re-
initiated from the beginning.

A temporary DOS file (TAR-
GET.TMP) is generated on the de-
fault drive by TARG1 to hold an im-
age of itself. TARG2 expects to find
this file on the default drive. If this
file is missing, TARG2 requests that
it be replaced ("TARGET.TMP not
found on default drive.").

A warning ("<address> <ID.>
—Code verb assumed.") is issued by
both TARG1 and TARG2 if there is a
verb with an address in its code field
that the recompiler cannot recognize.
The offending verb will be included
in the target machine code and the
recompiler will continue.

The most common problem we
have encountered while using this
target compiler has been caused by a
verb missing from the application
code that seemingly should have been
included. This situation arises be-
cause the missing verb is referenced
only indirectly (as a literal for in-
stance). That is, the missing verb's
pfa is compiled into a verb as a literal
by the verb ' at compile-time. If all
references to a verb are of this type
then the recompiler will not include
this verb in the application code. The
recompiler was designed this way be-
cause there is no way to anticipate
the intention of the application pro-
grammer for the literals compiled
into the definition of a verb.

If the application program crashes,
the programmer should first inspect
the source code to see if there are any
missing verbs. This can be done by
checking all verbs referenced by liter-
als (from ') against the verbs includ-
ed in the application code listed in the
DOS application code map file. If a
verb is missing, then the application
code must be changed to force inclu-
sion of the missing verb by explicitly
referencing that verb.

It is our intention that verbs not re-
quired for the run-time activity of the
application program not be included
in the target machine code. We have
found that the outer interpreter verbs
or the disk block virtual memory verbs
are sometimes included in the code

when they have no intended use in the application. The most common source of this problem is run-time error handling of Forth through the verb <ABORT">. This verb leads the recompiler far into the outer interpreter and BLOCK verbs of the host Forth system. To avoid including unwanted verbs, the application programmer needs to identify an unwanted verb (e.g. QUIT or R/W etc.). A program has been devised (to be described elsewhere) to determine the threads that lead to the unwanted verbs. The application programmer can redefine the particular verb that leads to the unwanted verbs. Example #3 demonstrates this situation. In this example some of the verbs of the interpreter have been redefined to exclude any reference to the disk block virtual memory verbs. This is quite a general procedure and demonstrates the great flexibility of Forth in allowing procedures that are central to the functioning of the host system to be redefined

for inclusion in the target code.

Inside the Target Compiler

The verb TARGET sequentially invokes the major tasks required by the recompilation process. The first of these tasks, APPTAG, tags (sets bit 5 of the header, the SMUDGE bit) all verbs in the host that are to be included in the application code.

The constant HEADERS, if it is non-zero (that is, the cfa of the application root verb), signals that all verbs necessary for the functioning of both the application root verb and the cold start verb (referenced by the constant TCOLDCFA) are to be tagged. In this case, headers will be included in the target code. The cold start routine must lead to an interpreter (such as QUIT).

If the constant HEADERS is zero, then only those verbs needed for the functioning of the cold start verb are tagged and the cold start verb must thread to the application program. In

this case, headers will not be included in the target code. If the interpreter or other unwanted verbs are unintentionally tagged, the source code must be modified to exclude references to unwanted verbs.

To search the tree structure defined by the application program, the recursive verb TREETAG is initiated with the cfa of the cold start of the application on the stack. TREETAG determines the type of the verb being examined at that node in the tree. Variables, constants, and machine-code verbs are simply tagged and TREETAG terminates at that level of recursion. If TREETAG is examining a colon verb, then each entry in the parameter field is the cfa of another verb that must be searched. TREETAG puts a cfa on the stack and then calls itself. This process is repeated for all entries in the parameter field. The verb PFSKIPS handles entries in the parameter field that are exceptions (entries that are to be skipped, such as

branches and character strings).

Verbs that are "CREATE DOES>" defining verbs are tagged by having VERBTAG find their name field addresses with the verb MAXNFA. MAXNFA recursively searches the entire host dictionary (when initiated properly; see the glossary, page 69) and returns the maximum nfa less than the address of the DOES> part of the "CREATE DOES>" construction. Further searching of the DOES> portion of the defining verb is performed by the verb CREATE-DOES>. This verb functions similarly to TREETAG and recursively calls TREETAG by the forward-reference-executing verb >TTAG. At present, the CREATE portion of the CREATE-DOES> construction will be included (wastefully) in the target code; however, no searching or tagging of the verbs referenced in this portion is performed. The exact form of "CREATE DOES>" defining verbs varies with different Forth implementations; some adjustments to the recompiler may be necessary with other FIG-Forth model systems.

APPTAG directs TFIX;CODE and TFIXNULL to tag certain verbs that are required in the target code; also the root verb of the application is searched if headers are requested.

VERBLISTS directs the creation of two lists on the parameter stack. The first list (proceeding down in memory) is a numerically sorted list of all name field addresses in the host dictionary. This list reflects the physical order of the verbs in memory rather than the order defined by the links within headers of the verbs. This physical ordering is useful during the elimination of unwanted verbs. The second verb list contains pairs of numbers. There is one pair for each tagged verb in the host. The first number is the address of the beginning of the memory segment for that tagged verb in the host system. This is either a cfa or an nfa, depending on whether or not headers are to be included in the target code. The second number of each pair is the numerical difference between the first number of the pair and the address of the start of that verb in the target ma-

chine code. This information will be used to recalculate all of the addresses in the target code to reflect the squeezing out of the unnecessary verbs from the host Forth system.

The verb FRAMEIN sets aside the space (see Figure 2, page 57) on the stack for the nfa list and the target offsets list. The size of this space is determined by the verb COUNT-VERBS which recursively searches all vocabulary branches of the host dictionary, determining the total number of verbs in the host and how many of those verbs are tagged (the variables VERBS# and TVERBS#). FRAMEIN

initializes certain pointers to these lists and also initializes the beginning and ending entries that are used as signals for the list searching verbs.

The verb OFFSETS enters into the target offsets list the starting address of the verbs to be in the target code. A running count of the verb locations to be in the target code is calculated. This number is used to calculate the difference between the host address and the target code address of each verb. For each verb that will be in the target, the host code address, paired with the offset (the difference calculated above), is entered into the tar-

get offsets list.

The name field addresses of all verbs in the host are entered into the nfa list by the verb NFALISTFILL. NFALISTFILL fills this list in the sequence followed by the recursive-branch-searching routine when searching the host dictionary (starting at the verb pointed to by FORTH CONTEXT @ @). This list is not necessarily in the same sequence as the physical order of the dictionary, and thus must be sorted. Since the list usually has large areas that are already in sequence with just the order of these blocks out-of-sequence, the sort can be performed rapidly by the verb WINDOWSORT. This routine detects the limits of the out-of-sequence blocks with WINDOWTOP, FRAMEBOTTOM, and WINDOWBOTTOM. The blocks are then moved into sequence by OPENMOVE.

The relocation routines compare the host Forth image generated by TARG1 (in the DOS file TARGET.TMP) with the current resident image (TARG2). These images should differ only for addresses, and the addresses should be different by exactly 201H. All addresses within the verbs being included in the target machine code must be recalculated to account for the squeezing out of unwanted verbs in the host. Certain addresses at the beginning of the target code must also be patched to reflect the memory map of the target machine.

Either of two pairs of routines direct this relocation operation, depending on whether headers are to be included in the target code (either RELOCATE and VERB>TARG, for headers, or RELOCATENH and VERB>TARGNH, for no headers). RELOCATE directs DISCARD to advance the file containing the image of TARG1 past areas of the file that are to be discarded. The verb VERB> TARG then copies the verb currently being examined by RELOCATE to the target image file. All addresses detected by VERB>TARG are examined by REPAIRADD, which returns the address reflecting the relocation of the current verb. If a link address is being adjusted, REPAIRADD directs NEWLINK to determine a new link address to account for intervening verbs in the host that are not to be included in the target code. In this process, the host dictionary structure is lost and the target dictionary links will reflect the physical order of its verbs. REPAIRADD detects whether the addresses need to be patched by PATCHADD to take into account the memory map of the target machine. REPAIRADD then consults the target offsets list to adjust the addresses to reflect the relocation of code.

If numbers in the two images are not equal, the verb VAR examines the code to determine if the number is in the data area of a variable. If this is not true, an error is issued by VAR and a dump produced by DUMP/=. In this case the tagged verbs are all untagged to allow examination of the host system from the outer interpreter by the application programmer.

If no headers were to be included, the same steps as above occur except that RELOCATENH and VERB> TARGNH are used. Because headers are not to be included in the target code, the vocabulary links do not need to be recalculated.

After all tagged verbs have been relocated and the application code file closed, a map of the memory allocation of the target machine is displayed by TARGETMEMORY. During the relocation process, each verb to be included in the target code is displayed on the screen along with its address in the target machine. This information is also copied to a file if requested (1 ' MAPFILE !) by the application programmer.

The DOS file interface is identical for both the CP/M-80 and MSDOS versions. The primary verbs are FC@ and FC!. These verbs fetch or store a byte from the sequential DOS file specified by the file specification address on the stack. The defining verb FILE creates a verb that functions like a verb created by VARIABLE. The verb created by the defining FILE has three parts: first, a count index byte; second, a disk memory area (128 bytes); and third, a DOS filename specification area (file control block). When the verb is executed at run-time, the address of the count byte is left on the stack. The verb DMA will advance this address to the disk memory area, and the verb FCB will advance the count byte address to the DOS filename specification area (file control block). The verbs FILE@ and FILE! are used by FC@ and FC! to fetch or store the next sequential segment to or from the disk memory area. Note that the file areas can be reused as often as desired. The verbs SPACEFILE, SPACESFILE, CRFILE, TYPEFILE and U.RFILE perform output functions to a file that are the same as analogous screen display verbs.

There are other verbs involved in maintaining the DOS file directory and in opening and closing DOS files. These verbs all return status flags except SETDMA. The verb FILENAME? queries a DOS filename from the keyboard and then sets up the filename specification area to be ready to search for or open a DOS file. File specification can also be loaded by the string-handling verb PUT$ with string verbs created by the defining verb STRING.

Complex Forth Constructions

Forth allows great flexibility in generating new programming constructions. Rather than trying to anticipate all possible new constructions, we decided to have the recompiler handle simple cases, but to allow the application programmer easy access to change the recompiler to handle more complex cases. Examples of cases that the recompiler does not currentiy handle are: 1.) a new verb with the function of colon; 2.) verbs that have an unusual effect on the parameter field of a verb being compiled (such as branches and literals).

The recompiler does properly handle simple CREATE-DOES> constructions (such as the defining verb STRING, screen #125). A verb such as F$ (screen #128) would be included in the application code if it had been referenced. The application code would also include the defining verb (STRING) since this verb contains the run-time action of the defined verb (F$). All other verbs necessary for the function of the DOES>

portion of the defining verb would also be included.

If the recompiler finds a value in a code field of a verb that it does not recognize, then that verb will be included in the application code; however, no further searching of this verb occurs. If this verb was the result of a new colon-type definition, then not searching this verb may result in some verbs that are needed not being included in the application code.

The recompiler may not generate the proper application code if there are bytes in the parameter field of a colon-defined verb that are not a code field address (or special instances that the recompiler has been instructed to search for, for example, branches and literals).

The recompiler has been designed to allow cases such as those described above to be easily included in the recompiler by the application programmer. The verb PFSKIPS on screen #137 must be augmented to handle special bytes within the parameter field of colon-defined verbs. The verb TREETAG on screen #139 can be augmented to handle cases such as newly defined colon-type definitions.

The verb EXIT must not be used since the verb tagging routines assume that the cfa of the verb EXIT terminates each colon defined verb. The cfa of EXIT must be the last cfa in the parameter field of a colon verb and must not appear anywhere else in the definition.

Availability of the Ready-to-Run Target Compiler from CICFIG

This project grew out of the needs of several of the CICFIG members for a target compiler to use to generate PROMable code for our work on various instrumentation control applications. We are placing the result of our efforts in the public domain because we believe there is a broad-based need for such a package. This target compiler fulfills a large part of our initial desires. It has not been ported to as large a variety of microprocessors as we had initially hoped, but this may change after the release of this code.

The address relocation process is not as fast as we had hoped; however, the modular design of the recompiler will allow us easily to replace these routines with more efficient ones in a later release.

Since the generation of TARG1. COM and TARG2.COM may be a bit complicated due to the recoding of certain parts of the kernel of the host Forth system, we have decided to distribute (at a cost of $30) a package containing the target compiler and all the parts of the target compiler generator. This package contains: the ready-to-run target compiler (TARG1 and TARG2), the verb tag tracer, and a program to filter the CP/M-80 target code to move the return stack pointer to RAM. Also included are the sources (Forth blocks and DOS files) for the recompiler, the examples, and the kernels. Documentation on the generation of the target

compiler and its use to target-compile application code are also included.

At present we are supplying this material for use on MSDOS compatible computers and on CP/M-80 compatible systems. Other systems may become available later and we should be contacted. Disk formats which are available are MSDOS Version 2 IBM-PC DSDD 5¼-inch 8-sector, CP/M-80 Version 2.2 standard 8-inch and CP/M-80 NorthStar Horizon SSDD 5¼-inch.

Address of CICFIG

CICFIG
Dept. of Electrical and Computer Engineering
University of Illinois
1406 W. Green St.
Urbana, Illinois 61801

Acknowledgements

Glossary for the CICFIG Recompiler Source Code

The following abbreviations are used in the stack pictures.

a   address (16 bit addressing)
b   byte
cfa   code field address
f   flag (0 = false)
fa   file address (the address of the first byte of a data area created by the defining word FILE)
n   number (16 bit signed)
nfa   name field address
pfa   parameter field address

!FIX (--- a) scr# 122 A variable used to avoid patching the link field of the verb ! when headers are to be included in the target code.

0BRANCHCFA (--- cfa) scr# 124 A constant that leaves the cfa of 0BRANCH.

;S (---) scr# 126 Put a zero (null) at the next byte of the input stream to terminate interpretation of a screen.

<+LOOP>CFA (--- cfa) scr# 124 A constant that leaves the cfa of <+LOOP>.

<."">CFA (--- cfa) scr# 124 A constant that leaves the cfa of <.">.

</LOOP>CFA (--- cfa) scr# 124 A constant that leaves the cfa of </LOOP>.

<;CODE>CFA (--- cfa) scr# 124 A constant that leaves the cfa of <;CODE>.

<ABORT">CFA (--- cfa) scr# 124 A constant that leaves the cfa of <ABORT">.

<LOOP>CFA (--- cfa) scr# 124 A constant that leaves the cfa of <LOOP>.

>TTAG (---) scr# 136 Execute the verb TREETAG. This allows a forward reference to TREETAG during compilation of the recompiler.

APPTAG (---) scr# 140 This verb routes the verb tagging routines.

ASCII (--- b) scr# 127 Leave the number that is the ASCII code for the next character after space in the input stream.

BRANCHCFA (--- cfa) scr# 124 A constant that leaves the cfa of BRANCH.

CASE (---) scr# 127 Defined in "Eaker Case Statement" *FORTH Dimensions* II/3 p 37, 9/80.

CLEANUPDOS (---) scr# 155 Flushes and closes the DOS files used by the relocation routines.

COUNTVERBS (nfa ---) scr# 134 Search the dictionary of the host Forth system from the verb with name field address nfa to the beginning of the dictionary. Keep a running count of the number of verbs encountered in the variable VERBS#, and the number of tagged verbs in the variable TVERBS#. This verb is recursive and is started by leaving the nfa obtained from FORTH CONTEXT @ @. The entire host Forth system is searched through recursive calls to COUNTVERBS to include all vocabularies.

CREATE-DOES> (a ---) scr# 138

Tag the verb with the largest nfa less than the address a. If a points to the DOES> portion of a verb that is a CREATE-DOES> construction, continue to TREETAG all verbs which the DOES> portion of this verb requires. Otherwise, indicate an exception ("--- Code verb assumed") and continue without searching the verb that the address a points into.

CRFILE (fa ---) scr# 134 Append a <cr> to the DOS file specified by the file address fa.

DDRIVE (---) scr# 141 Prompt the operator, and pause until a key is pressed.

DISCARD (n ---) scr# 152 Discard n bytes from the file IMG (TARGET.TMP).

DJMPLEN (--- n) scr# 120 An implementation specific constant that leaves the number of bytes of the machine code sequence that is a jump to DODOES at the start of the DOES> portion of a "CREATE DOES>" defining verb.

DMA (fa --- a) scr# 128 Leave the address of the disk transfer area of the file specified by the file address fa.

DOCODE (--- a) scr# 122 A variable used in TREETAG to determine if the verb being currently examined is a code level verb.

DOCOL (--- a) scr# 124 A constant that leaves the address of DOCOL.

DOCON (--- a) scr# 124 A constant that leaves the address of DOCON.

DOESJMP (--- b) scr# 120 An implementation specific constant that leaves the first byte of the machine code sequence that is a jump to DODOES at the start of the DOES> portion of a "CREATE DOES>" defining verb.

DOUSE (--- a) scr# 124 A constant that leaves the address of DOUSE.

DOVAR (--- a) scr# 124 A constant that leaves the address of DOVAR.

DOVAR? (--- a) scr# 122 A variable used as a flag by the code relocation routines to signal that a variable is currently being relocated.

DOVOC (--- a) scr# 124 A constant that leaves the address of DOVOC.

DOWN (--- a) scr# 122 A variable whose value is used as a pointer to within the name field address list.

DUMP (a n ---) (defined in MVP-FORTH Utilities) Display the contents of the n bytes of memory starting at address a.

DUMP/= (a ---) scr# 148 Display the contents of memory in the vicinity of the address a. This will help to identify the section of the Forth system that was not equivalent to the image in the DOS file TARGET.TMP. In addition, the DOS files are closed, all tagged verbs are untagged, and then the program aborts.

ENDCASE (---) scr# 127 See CASE.

ENDOF (---) scr# 127 See CASE.

EXITCFA (--- cfa) scr# 124 A constant that leaves the cfa of the verb EXIT.

**F$** (--- a n) scr# 128 A string created by the defining verb STRING. There are n bytes currently in F$, starting at address a. This string holds the name of the DOS file TARGET.TMP.

FC! (b fa ---) scr# 133 Append the byte b to the file specified by the file address fa.

FC@ (fa --- b) scr# 133 Get the next byte b from the file specified by the file address fa.

FCB (fa --- a) scr# 128 Leave on the stack the address of the file control block area specified by the file address fa.

FCLOSE (fa --- f) scr# 131 Close the DOS file specified by the file address fa. Leave the flag FFFF hex if the file cannot be found, 0 if successful.

FCREATE (fa --- f) scr# 130 Make (and open) a new DOS file with the name specified by the contents of the FCB area of the file specified by the file address fa. Leave FFFF hex if no more directory space is available, 0 if successful.

FDELETE (fa --- f) scr# 130 Delete the DOS file with the name specified by the contents of the FCB area of the file specified by the file address fa. Leave FFFF hex if the file was not found, 0 is successful.

FILE (---) scr# 128 A defining word that creates a verb that func-

tions like a variable with 162 bytes. Use like VARIABLE: FILE <filename>. In the parameter field of <filename>, the first byte is used to keep a count on the disk memory transfer area from DOS (the 128 bytes where data transfers from DOS are stored), and the last 33 bytes are used as the DOS file control block for the DOS filename.

FILE! (fa --- f) scr# 132 Append the 128 bytes in the file transfer area of the file specified by the file address fa. Leave FFFF hex if the disk was full, 0 if the write was successful.

FILE@ (fa --- f) scr# 132 Read the next 128 bytes into the DMA area from the DOS file named in the FCB area of the file specified by the file address fa. Leave FFFF hex if no more data can be read, 0 if the read was successful.

FILECREATE (fa --- f) scr# 132 Make (and open) a new DOS file deleting if necessary an existing file with the same name. The DOS file is named by the contents of the FCB area of the file specified by the file address fa. Leave FFFF hex if no directory space was available, 0 if successful.

FILEEOF (fa --- f) scr# 132 Leave FFFF hex if the last allocated byte of the DOS file specified by the file address fa has been read; 0 otherwise.

FILENAME? (fa ---) scr# 129 Prompt the keyboard for a DOS filename. The filename is loaded into the FCB area of the file specified by the file address fa.

FINDTAG (a1 --- a2) scr# 146 Search the name field address list starting at address a1 and leave the the address a2 of the next location in the list that has a value that points to a tagged verb.

FIXLINK? (--- a) scr# 122 A variable used by the relocation routines when the target code is to contain headers. A non-zero value indicates that the next encountered address is a link and must be patched to correct for the possible absence in the target code of the verb to which the link used to point to.

FIXSMUDGE (---) scr# 147 This verb sets to zero the smudge bit of all verbs which have been tagged by the tagging routines.

FLIP (n1 --- n2) scr# 126 Swap the high and low bytes of the word on the stack.

FNSEARCH (fa ---) scr# 130 Have DOS search for the next file that matches the file named (wildcards are usually used) in the FCB area specified by the file address fa. The matching DOS filename is returned in the disk memory transfer area last assigned by the verb SETDMA. Leaves FFFF hex if no match is found, 0 if a match is found.

FOPEN (fa --- f) scr# 131 Open a DOS file with the name contained in the FCB area of the file specified by the file address fa. Leaves FFFF hex if the file cannot be found, 0 if the file was successfully found.

FORTHORG (--- a) scr# 120 An implementation specific constant that leaves the address of the origin of the current host Forth system.

FRAMEBOTTOM (a1 n --- a2) scr# 144 Search down the name field address list from the pointer a1 until the value in the list is less than the number n. Leave the pointer a2 for this entry.

FRAMEIN (--- n1...nj) scr# 142 Construct the framework for the verb lists on the stack by leaving zeros for the j cells required. Certain locations within the lists and certain pointers to the lists are initialized. The verb RAZE will return the stack to its depth prior to FRAMEIN. The number j is calculated from the variables VERBS# and TVERBS#.

FREAD (fa --- f) scr# 131 Read the next 128 bytes from the file specified by the file address fa. These bytes are transferred to the disk transfer area specified by the last call to the operating system (SETDMA). Leave FFFF hex if there is no more data available, 0 if the read was successful.

FSEARCH (fa --- f) scr# 130 Have DOS search for the file named in the FCB area of the file specified by the file address fa. Leave FFFF hex if no match is found, 0 otherwise.

FWRITE (fa --- f) scr# 131 Append 128 bytes from the last disk transfer area set (SETDMA), to the file named in the FCB area specified the file address fa. Leave FFFF hex if the disk was full, 0 if the write was successful.

HEADERS (--- n) scr# 123 A constant that leaves either 0 if the recom-

piler is to generate headerless code, or the pfa of the root verb of the application program if headers are to be included in the target code.

HITKEY (---) scr# 141 Pause until a key is depressed on the console.

ID.FILE (nfa fa ---) scr# 135 Copy a verb name to the file specified by the file address fa.

IMAGESAVE (---) scr# 141 Copy an image of the current Forth system from FORTHORG to HERE to the DOS file TARGET.TMP.

IMG (--- fa) scr# 128 A file specifier created by the defining verb FILE. Leaves the address fa of the disk transfer area count byte.

ISFF (f1 --- f2) scr# 130 Leaves FFFF hex if f1 is FF hex, 0 otherwise.

ISZERO (f1 --- f2) scr# 130 Leaves 0 if f1 is 0, FFFF hex otherwise.

LB (n --- b) scr# 128 Leave the low byte of the number n.

LITCFA (--- cfa) scr# 124 Leave the cfa of the verb LIT.

LOADFCB (a1 n a2 ---) scr# 141 Zero 33 bytes starting at address a2 (an FCB area), and move n bytes

from address a1 to address a2+1.

LTIB (--- a) scr# 120 A constant that leaves the address of a cell in low memory that contains a pointer to the terminal-input-buffer.

**M$** (--- a n) scr# 128 A string created by the defining verb STRING. There are n bytes currently in M$, starting at address a. This string holds the DOS filename attribute MAP.

MAP (--- fa) scr# 128 A file specifier created by the defining verb FILE. Leaves the address of the disk transfer area count byte.

MAPFILE (--- f) scr# 123 A constant that leaves a zero if no map file is to be generated; otherwise, signal the generation of a DOS ASCII file <filename>.MAP with the names and addresses of the verbs included in the target code.

MAXNFA (nfa1 a nfa2 --- nfa3 a) scr# 135 This verb recursively searches the entire host dictionary starting at nfa2 (FORTH CONTEXT @ @) and returns nfa3 which is the maximum nfa that is less than or equal to a. When started, nfa3

should be 0.

NEWLINK (nfa1 --- nfa2) scr# 148 Find the host name field address nfa2 of the verb that when moved to the target code, should be pointed to by the link field of the verb with the host name field address nfa1.

NEXTNFA (nfa1 --- nfa2) scr# 150 Leave the name field address nfa2 of the next verb in physical order in the host Forth system after the verb with the name field address nfa1. If the verb with name field address nfa1 is the last verb in the dictionary (in physical order), then leave HERE.

NFALISTFILL (nfa ---) scr# 143 Search the dictionary of the host Forth system from the verb with name field address nfa, to the beginning of the dictionary. Enter the name field address of all verbs encountered in the name field address list. This verb is recursive and is started by leaving the nfa obtained from FORTH CONTEXT @ @. The entire host Forth system is searched through recursive calls to NFALISTFILL to include all vocabularies.

NFTAG (nfa ---) scr# 136 Set the smudge bit of the verb with name field address nfa to 1.

NLIST (--- a) scr# 123 A constant that leaves the address of the first entry in the name field address list.

OF (---) scr# 127 See CASE.

OFFSETS (---) scr# 146 Enter into the tagged verb offsets list the name field addresses (or code field addresses) of all tagged verbs in the host Forth system. With each entry, store the difference between this address in host Forth system and the address that verb will be at in the target code.

OPENMOVE (a1 a2 a3 ---) scr# 145 Insert at list pointer a2, the section of the name field address list that starts at pointer a1 and ends at pointer a3. See WINDOWSORT.

PATCHADD (a1 --- a2 f) scr# 149 If the flag f is 1, then the address a2 found at the address a1 should be corrected to a new address for the target code. If the flag f is 0, then the address a2 should not be corrected. This is used to allow loading of target code RAM addresses.

PATCHMAX (--- a) scr# 120 An implementation specific constant

that leaves the address above which in the host Forth system no addresses will need to be patched by PATCH-ADD for inclusion in the target code.

PFSKIPS (a1 --- a1 or a2 0) scr# 137 Check to see if the cfa (which is an entry in the parameter field of a colon verb) pointed to by a1 is in need of special handling, such as branches, loops and text. Leave the address a1 if this cfa is not an exception. Leave the address a2 and a 0 flag to bypass searching the verb referenced by a1, and continue searching at the pointer a2 + 2.

PUT$ (a n ---) scr# 125 Move the bytes (string) in the input stream starting after the next space character to the next $ character (exclusive) to the address a. The byte before the address a stores the current number of bytes in the string and the byte below that holds the maximum number of bytes in the string (see STRING). QUIT if the string is too long. The number n is ignored.

RAZE (n1...nj ---) scr# 142 Remove the verb lists (j cells) from the parameter stack (see FRAMEIN). The number j is calculated from the variables VERBS# and TVERBS#.

RELOCATE (--- n) scr# 152 Used when headers are to be included in the target code. This word organizes the relocation of the necessary sections of the host Forth system to generate the target code. The total number of bytes n in the target code is left on the stack.

RELOCATENH (--- n) scr# 154 Used when headers are not to be included in the target code. This word organizes the relocation of the necessary sections of the host Forth system to generate the target code. The total number of bytes n in the target code is left on the stack.

REPAIRADD (a1 --- a2) scr# 150 Leave the address a2 which is the value at the host Forth address a1, corrected to reflect the address relocation of the code to be included in the target code.

ROOF (--- a) scr# 123 A constant that leaves the address of last name field address in the name field address list.

SETDMA (fa ---) scr# 131 Instruct DOS that the disk transfer area is at memory location fa + 1.

SETUPDOS (---) scr# 155 Setup the DOS files for the relocation routines of the recompiler.

SFLAG (f1 n --- f2 n) scr# 136 This verb changes the flag f1 to f2 (f2=f1−1), This is used to determine the none-of-the-above case of the CASE construction in the verb TREETAG.

SPACEFILE (fa ---) scr# 134 Append a space character to the file specified by the file address fa.

SPACESFILE (n fa ---) scr# 134 Append n space characters to the file specified by the file address fa.

STRING (n ---) scr# 125 A defining word used to create a dictionary entry for the verb with the name of the next character string in the input stream. The new verb will be allotted n + 2 bytes. The first byte is the maximum allotted space for the string, the next byte is the current length of the string. At run-time the new verb will leave on the stack the address of the first byte of the string and the number of bytes in the string.

T'STREAM (--- a) scr# 129 Functions as 'STREAM except that the verb BLOCK has been removed

T#BUFF (--- n) scr# 123 A constant that leaves the number of disk block buffers required by the application program.

TARG (--- fa) scr# 128 A file specifier created by the defining verb FILE. Leaves the address of the disk transfer area count byte.

TARGET (pfa or 0 ---) scr# 157 This verb routes the operation of the target compiler. If the top of the stack has a zero, then no headers will be included in the target code; otherwise, this value is the parameter field address of the root verb of the application program and headers will also be included in the target code.

TARGETMEMORY (n ---) scr# 156 Display a memory map of the target code that has n bytes.

TARGMAP (n nfa ---) scr# 148 Display the verb name with name field address nfa and display the target code address (n + TARGORG) of that verb. Copy this same information to a DOS file if a map file has

been requested.

TARGORG (--- a) scr# 123 A constant that leaves the address of the origin of the target code.

TBUF1 (--- a) scr# 123 A constant that leaves the address of the start of the disk block buffer area of the target code.

TCOLD1 (--- a) scr# 120 An implementation specific constant that leaves the address of the cell in low memory that contains the address of the first verb to be executed upon start-up of the target code.

TCOLDCFA (--- cfa) scr# 123 A constant that leaves the cfa of the first verb to be executed upon start-up of the target code.

TEM (--- a) scr# 123 A constant that leaves the address of the top of the disk block buffer area in the target code.

TFIX;CODE (---) scr# 140 This verb tags certain ;CODE verbs that are mandatory in the target code.

TFIXNULL (---) scr# 140 This verb tags the verb X (null) and then patches the header for this verb.

TLIST (--- a) scr# 123 A constant that leaves the address of the start of the target verb offset list.

TPICK (n1 --- n2) scr# 126 Functions as PICK except that ABORT" has been removed.

TR0 (--- a) scr# 123 A constant that leaves the address of the start of the target code return stack pointer for initialization of the inner-interpreter of the target machine.

TROLL (n ---) scr# 126 Functions as ROLL except that ABORT" has been removed.

TREETAG (cfa ---) scr# 139 This recursive verb tags (sets the smudge bit to 1) all verbs in the host Forth system that are required for the function of the verb whose code field address is cfa.

TREETAGCFA (--- cfa) scr# 123 A constant that leaves the cfa of TREETAG. Used to allow forward referencing at compilation time within TARGET.

TSP0 (--- a) scr# 123 A constant that leaves the address of the start of the target code parameter stack pointer for initialization of the inner-interpreter of the target machine.

TVARIABLE (---) scr# 126 A de-

fining word used in the form: TVARIABLE <varname>. When TVARIABLE is executed, it creates the verb <varname> in the host dictionary with its parameter field initialized to the next allotted cell in the variable RAM area of the target machine. This allows variables to be implemented in RAM when the target code is in ROM. The application programmer must set up the target variable RAM area by loading the constant TVRAM.

TVERBS# (--- a) scr# 122 A variable used to hold the number of verbs that have been tagged by the tagging routines.

TVRALLOT (n ---) scr# 126 Add the number n to the variable TVRAMOFF. This allows target variable RAM bytes to be allotted.

TVRAM (--- a) scr# 123 A constant that leaves the address of the start of the target variable RAM area. This value must be loaded by the application programmer if needed.

TVRAMOFF (--- a) scr# 122 A variable that contains the current number of bytes of the target variable RAM area that have been allotted.

TWORD (b --- a) scr# 129 Functions as <WORD> except that ABORT" has been removed.

TYPEFILE (a n fa ---) scr# 134 Copy n characters starting at address a to the file specified by the file address fa.

U.RFILE (n1 n2 fa ---) scr# 134 Convert the number n1 to an unsigned ASCII number right justified n2 spaces. Append these ASCII characters to the file specified by the file address fa.

U8 (--- a) scr# 120 A constant that leaves the address of the cell in low memory that contains the value to initialize the eight user variable.

VAR (a ---) scr# 149 If the address a is the pfa of a VARIABLE defined verb (or the return stack pointer cell address) then copy the contents of that cell to the DOS file specified by TARG. Otherwise, dump the contents of the host Forth system in vicinity of the address a and indicate an error.

VERB>TARG (nfa ---) scr# 151 The host Forth verb (including the

header) with name field address nfa is copied to the DOS file specified by TARG. Address relocation and patches are performed.

VERB>TARGNH (cfa ---) scr# 153 The host Forth verb (excluding the header) with code field address cfa is copied to the DOS file specified by TARG. Address relocation and patches are performed.

VERBLISTS (---) scr# 147 This routine routes the construction and filling in of the verb lists on the parameter stack.

VERBS# (--- a) scr# 122 A variable used to hold the number of verbs that are in the host Forth system dictionary.

VERBTAG (a ---) scr# 136 Tag (set the smudge bit to 1) the verb whose code field address or DOES> jump address is a.

WINDOWBOTTOM (a1 a2 --- a3) scr# 144 Examine the values in the name field address list starting with the element pointed to by a1. Continue so long as they are decreasing and greater than the contents of the element of the list pointed to by a2. When this is not true, leave a3, which points to the element just beyond this point.

WINDOWSORT (a ---) scr# 145 Sort the list of unsigned numbers (name field addresses) between the high memory address a and low memory location specified by the value FFFF hex in the list. At the end of the sort, change the entry at the bottom of the list to FORTHORG.

WINDOWTOP (a1 --- a2) scr# 144 Index down the list of unsigned numbers starting at pointer a1 until the values in the list are no longer monotonically decreasing. Leave pointer a2 to where this occurs. See WINDOWSORT.

X (---) scr# 140 This is a pseudonym for null that does not have the compile-time parts of the host system Forth verb null. This version of null will be present in the target code.

Procedure

TARG1

    (A) load screen with memory map   (from TARG1.COM DOS)
    (B) load screen(s) with application   (scr# 200, example #1)
    (C) load screen with cold start   (scr# 201, example #1)
                                          (scr# 201, example #1)

0 TARGET

    (control will be returned to the operating system)   (0 => no headers)

TARG2

                                          (from TARG2. COM DOS)

    (A-C) exactly as above   (200 201 THRU)

0 TARGET

                                          (0 => no headers)

    (upon request, supply the DOS filename for the target code)
    (control will be returned to the interpreter when complete)

Forth Compiler Listing (Text begins on page 52)

```
SCR #200
 1 ( No headers, No block verbs; Target Memory Map Constants) HEX
 2 ( This program outputs to printer, ASCII chars typed in in hex)
 3
 4 1 ' MAPFILE !
 5
 6 6000 ' TVRAM !
 7 6000 ' TEM !
 8 TEM ' TBUF1 !
 9 TBUF1 52 - ' TRO !
10 TRO A0 - ' TSPO !
11 100 ' TARGORG !
12 DECIMAL
13
14
15
16


SCR #201
 1 ( LP TINPUT APP TCOLD) HEX
 2
 3 : LP  ( b --- )  5 SWAP SYSCALL ;
 4
 5 : TINPUT ( --- n )  0 0 ." ? " QUERY 1 TWORD CONVERT DROP DROP ;
 6
 7 : APP  ( --- ) HEX
 8   ." To send an ASCII character to the printer," CR
 9   ." type a HEX number and (cr) after '?' {FF to terminate}." CR
10 BEGIN TINPUT DUP FF = NOT WHILE SPACE LP REPEAT 0 0 SYSCALL ;
11
12 : TCOLD ( --- ) 0 EPRINT !
13 INIT-USER UP @ 6 + 30 CMOVE PAGE APP ;
14
15 ' TCOLD CFA ' TCOLDCFA !
16 DECIMAL


SCR #202
 1 ( No headers, No block verbs; MINIMUM Test; Target Memory Map)
 2 HEX
 3 0 ' MAPFILE !
 4
 5 6000 ' TVRAM !
 6 6000 ' TEM !
 7 TEM ' TBUF1 !
 8 TBUF1 ' TRO !
 9 TRO A - ' TSPO !
```

```
10 100 ' TARGORG !
11 DECIMAL
12
13
14
15
16


SCR #203
 1 ( TCOLD) HEX
 2
 3 : TCOLD  0 0 0 0 21 INTCALL ;
 4
 5 ' TCOLD CFA ' TCOLDCFA !
 6 DECIMAL
 7
 8
 9
10
11
12
13
14
15
16


SCR #204
 1 ( Headers, No block verbs; Target Memory Map Constants) HEX
 2
 3 1 ' MAPFILE !
 4
 5 6000 ' TVRAM !
 6 6000 ' TEM !
 7 TEM ' TBUF1 !
 8 TBUF1 52 - ' TRO !
 9 TRO A0 - ' TSPO !
10 100 ' TARGORG !
11 DECIMAL
12
13
14
15
16
```

DPMA Houston '85,

WHERE COMPUTERS GET DOWN TO BUSINESS

Data Processing Management Association's
Annual International Computer Conference and Business Exposition

Oct. 28-30, 1985

Albert Thomas Convention & Exhibit Center
Houston, Texas

**Open To All In The Computer or Related Industries
. . .Seminars. . .Workshops. . .Panel Discussions
. . .Tutorials. . .Half-Day Sessions. . .Exhibits**

Conference Highlights

- Impact on Information Management
- Small Systems Environment
- Controlling the Revolution
- Back to Basics
- And many more topics
- Conference On-Site Registration Begins on October 27. All Computer Industry Professionals are Welcome to Attend.

Business Exposition

Products, services and supplies by the nation's leading companies including AT&T. . .Digital Equipment Corp.. . .Eastman Kodak. . .IPM. . .Radio Shack. . .Xerox. . .Cincom Systems. . .GE. . .Lanier. . .and many others.

Houston's the perfect site!

As the nation's fourth largest city, Houston has immense financial wealth with expanding opportunities and growth. It also has one of the few remaining center-city areas that is still fun, exciting and comfortable—day and night. The Albert Thomas Convention and Exhibit Center is located in the middle of thriving downtown Houston.

For complete details on DPMA Houston '85, fill out the coupon below.

DPMA HOUSTON '85
WHERE COMPUTERS GET DOWN TO BUSINESS
Oct. 28-30, 1985

**Send to DPMA Houston '85
505 Busse Highway
Park Ridge, IL 60068-3191**

Please send:
☐ Conference registration materials
☐ Free exhibit-viewing tickets
☐ My company's interested in exhibiting
☐ Membership information

Name/Title

Company

Address

City/State-Prov./Zip-Postal Code

Telephone

Forth Compiler Listing

```
SCR #205
 1 ( TABORTCFA TNABORT TNUMBER T-FIND)
 2
 3 0 CONSTANT TABORTCFA
 4
 5 : TNABORT ." NOT RECOGNIZED" TABORTCFA EXECUTE ;
 6
 7 : TNUMBER  ( a --- d )
 8 0 0 ROT DUP 1+ C@ ASCII - = DUP )R + -1 DPL !
 9 CONVERT DUP C@ BL )
10 IF DUP C@ ASCII . = NOT IF TNABORT THEN O DPL !
11    CONVERT DUP C@ BL > IF TNABORT THEN
12 THEN DROP R) IF NEGATE THEN ;
13
14 : T-FIND  BL TWORD CONTEXT @ @ <FIND> ;
15
16


SCR #206
 1 ( TINTERPRET TQUIT TABORT)
 2
 3 : TINTERPRET  ( --- )
 4 BEGIN T-FIND
 5    IF DROP CFA EXECUTE
 6    ELSE HERE TNUMBER DPL @ 1+ NOT IF DROP THEN
 7    THEN
 8 AGAIN ;
 9
10 : TQUIT  ( --- )
11 BEGIN CR RP! QUERY TINTERPRET ." OK" AGAIN ;
12
13 : TABORT  ( --- )
14 SP! [COMPILE] FORTH TQUIT ;
15
16 ' TABORT CFA ' TABORTCFA !


SCR #207
 1 ( APP TCOLD)
 2 HEX
 3 : APP CR ." HELLO FROM APP" BEGIN ?TERMINAL ?DUP UNTIL ;
 4
 5 : T.R  ( n1 n2 --- ) )R 0 SWAP OVER DUP D+- (# #S ROT SIGN #)
 6 R) OVER TYPE ;
 7
 8 : TCOLD  ( --- ) 0 EPRINT !
 9 INIT-USER UP @ 6 + 30 CMOVE
10 PAGE ." QUIT OK" CR
11 INIT-FORTH @ ' FORTH 2+ !
12 DECIMAL TABORT HEX T.R ;
13 DECIMAL
14
15 ' TCOLD CFA ' TCOLDCFA !
16


SCR #208
 1 ( Block verbs, No headers; BLOCKASC; Target Memory Map) HEX
 2 ( Generates DOS-ASCII files of Forth disk block areas)
 3
 4 1 ' MAPFILE !
 5
 6 2 ' T#BUFF !
 7
 8 6000 ' TVRAM !
 9 6000 ' TEM !
10 TEM 404 T#BUFF # - ' TBUF1 !
11 TBUF1 52 - ' TRO !
12 TRO A0 - ' TSP0 !
13 100 ' TARGORG !
```

```
14 DECIMAL
15
16


SCR #209
 1 ( CHARS BLOCK)FOUT) HEX
 2
 3 : CHARS  ( a --- n )
 4 DUP C/L 1-
 5 BEGIN DUP C@ 20 = OVER 4 TPICK ) AND
 6 WHILE 1-
 7 REPEAT SWAP - 1+ ;
 8
 9 : BLOCK)FOUT  ( ba fa --- )
10 SWAP DUP 400 + SWAP
11 DO I CHARS I + I DO I C@ OVER FC! 1 /LOOP
12    D OVER FC! A OVER FC! C/L
13 /LOOP DROP ;
14 DECIMAL
15
16


SCR #210
 1 ( FOUT TINPUT BLOCKASC) HEX
 2
 3 FILE FOUT
 4
 5 : TINPUT ( --- n )  0 0 ." ? " QUERY 1 TWORD CONVERT DROP DROP ;
 6
 7 : BLOCKASC  ( --- )
 8 CR ." Starting block # (DEC)" TINPUT
 9 ." # of blocks" TINPUT
10 ." DRIVE (0,1)" TINPUT 0 MAX 1 MIN IF DR1 ELSE DR0 THEN
11 CR ." DOS file out name" FOUT FILENAME? FOUT FILECREATE DROP
12 OVER + SWAP
13 DO I BLOCK FOUT BLOCK)FOUT LOOP
14 1A FOUT FC! FOUT FILE! DROP FOUT FCLOSE DROP
15 CR ." Task complete." CR BYE ;
16 DECIMAL


SCR #211
 1 ( TCOLD) HEX
 2
 3 : TCOLD  ( --- ) 0 EPRINT !
 4 INIT-USER UP @ 6 + 30 CMOVE PAGE
 5 TBUF1 ' FIRST ! TEM ' LIMIT ! T#BUFF ' #BUFF ! EMPTY-BUFFERS
 6 CR 1 DENSITY ! FIRST USE ! FIRST PREV !
 7 DR0 40 ' C/L !
 8 DECIMAL BLOCKASC ;
 9
10 ' TCOLD CFA ' TCOLDCFA !
11 DECIMAL
12
13
14
15
16


SCR #212
 1 ( (TR/W))
 2
 3 : (TR/W)  ( n --- ) USE @ )R ROT USE ! SWAP MAX-DRV 0
 4 DO I DR-DEN DENSITY ! DUP BPDRV - -1 )
 5    IF BPDRV - I 1+ MAX-DRV - IF BYE THEN
 6    ELSE I DRIVE ! LEAVE
 7    THEN
 8 LOOP SPBLK # SPBLK 0
 9 DO DDUP T&SCALC
```

BUSINESS REPLY MAIL

FIRST CLASS PERMIT #27346, PHILADELPHIA, PA.

POSTAGE WILL BE PAID BY ADDRESSEE

SOFTWARE TOOLS FOR ADVANCED PROGRAMMERS

Dr. Dobb's Journal

P.O. BOX 13851
PHILADELPHIA, PA 19101

```
10    IF SEC-READ ELSE SEC-WRITE THEN
11    1+ 1024 SPBLK / USE +!
12 LOOP DDROP R) USE ! ;
13
14 ' (TR/W) CFA 'R/W !
15
16


SCR #213
 1
 2
 3
 4
 5
 6
 7
 8
 9
10
11
12
13
14
15
16

SCR #214
 1 ( Block verbs, No headers; ASCBLOCK; Target Memory Map) HEX
 2 ( Generates Forth disk block areas form DOS-ASCII files)
```

```
 3
 4 1 ' MAPFILE !
 5
 6 2 ' T#BUFF !
 7
 8 6000 ' TVRAM !
 9 6000 ' TEM !
10 TEM 404 T#BUFF # - ' TBUF1 !
11 TBUF1 52 - ' TRO !
12 TRO AO - ' TSPO !
13 100 ' TARGORG !
14 DECIMAL
15
16


SCR #215
 1 ( CURBLOCK BCOUNT FIN BC! LINE)BLOCK) HEX
 2
 3 TVARIABLE CURBLOCK   TVARIABLE BCOUNT
 4
 5 FILE FIN
 6
 7 : BC! ( b --- ) BCOUNT DUP @ 400 =
 8 IF 1 CURBLOCK +! UPDATE CURBLOCK @ BUFFER DROP 0 OVER ! THEN
 9 DUP @ DUP 1+ ROT ! PREV @ 2+ + C! ;
10
11 : LINE)BLOCK ( b --- ) DUP D =
12 IF 0 ELSE 7F AND BC! 1
```

```
13     BEGIN FIN FC@ 7F AND DUP D = NOT
14     WHILE BC! 1+
15     REPEAT DROP FIN FC@ DROP C/L SWAP -
16  THEN BEGIN DUP WHILE 1- 20 BC! REPEAT DROP ; DECIMAL


SCR #216
 1 ( TYPEFILENAME TIMPUT ASCBLOCK) HEX
 2 : TYPEFILENAME ( fa --- )
 3   DUP FCB 1+ 8 TYPE ." ." FCB 9 + 3 TYPE ;
 4
 5 : TIMPUT ( --- n )  0 0 ." ? " QUERY 1 TWORD CONVERT DROP DROP ;
 6
 7 : ASCBLOCK  ( --- )
 8   CR ." DOS in file name" FIN FILENAME? FIN FOPEN
 9   IF CR FIN TYPEFILENAME ." can't be opened." BYE THEN
10   CR ." Starting block # (DEC)" TIMPUT ." Drive # (0,1)" TIMPUT
11   0 MAX 1 MIN IF DR1 ELSE DR0 THEN
12   DUP CURBLOCK ! 0 BCOUNT ! BUFFER DROP
13   BEGIN FIN FC@ DUP 1A = NOT WHILE LINE)BLOCK REPEAT DROP
14   BEGIN BCOUNT @ 400 ( WHILE 20 BC! REPEAT
15   UPDATE SAVE-BUFFERS FIN FCLOSE DROP
16   CR ." Task complete." CR BYE ; DECIMAL


SCR #217
 1 ( TCOLD) HEX
 2
 3 : TCOLD ( --- ) 0 EPRINT !
 4   INIT-USER UP @ 6 + 30 CMOVE PAGE
 5   TBUF1 ' FIRST ! TEM ' LIMIT ! T#BUFF ' #BUFF ! EMPTY-BUFFERS
 6   CR 1 DENSITY ! FIRST USE ! FIRST PREV !
 7   DR0 40 ' C/L !
 8   DECIMAL ASCBLOCK ;
 9
10  ' TCOLD CFA ' TCOLDCFA !
11 DECIMAL
12
13
14
15
16


SCR #120
 1 ( Implementation specific constants for CP/M) HEX
 2
 3 INIT-FORTH 12 - CONSTANT FORTHORG   ( Host origin)
 4 148 CONSTANT PATCHMAX   ( end of patch area)
 5 110 CONSTANT TCOLD1     ( address for patch for jump to cold)
 6 116 CONSTANT U8         ( address for patch for user#8)
 7 118 CONSTANT LTIB       ( address for patch for Term-in-buffer)
 8 CD  CONSTANT DOESJMP    ( first byte of run-time create-does))
 9 3   CONSTANT DJMPLEN    ( # of bytes for run-time c-d jmp)
10 DECIMAL
11
12
13
14
15
16


SCR #121
 1 ( MS-DOS Implementation specific addresses) HEX
 2
 3 INIT-FORTH 1B - ' FORTHORG !
 4 151 ' PATCHMAX !                        ( MS-DOS)
 5 119 ' TCOLD1 !                          ( MS-DOS)
```

```
 6 11F ' U8 !                             ( MS-DOS)
 7 121 ' LTIB !                           ( MS-DOS)
 8 88  ' DOESJMP !                        ( MS-DOS)
 9 5   ' DJMPLEN !                        ( MS-DOS)
10 DECIMAL
11
12
13
14
15
16


SCR #122
 1 ( VARIABLES) HEX
 2
 3 VARIABLE TVRAMOFF  0 TVRAMOFF !
 4 VARIABLE DOCODE    0 DOCODE !
 5 VARIABLE VERBS#    0 VERBS# !
 6 VARIABLE TVERBS#   0 TVERBS# !
 7 VARIABLE DOWN      0 DOWN !
 8 VARIABLE FIXLINK? 0 FIXLINK? !
 9 VARIABLE !FIX      0 !FIX !
10 VARIABLE DOVAR?   0 DOVAR? !
11 DECIMAL
12
13
14
15
16


SCR #123
 1 ( CONSTANTS see Examples and Glossary) HEX
 2 0 CONSTANT MAPFILE      ( These constants are loaded by the user)
 3 0 CONSTANT T#BUFF
 4 0 CONSTANT TVRAM
 5 0 CONSTANT TEM
 6 0 CONSTANT TBUF1
 7 0 CONSTANT TR0
 8 0 CONSTANT TSP0
 9 0 CONSTANT TARGORG
10 0 CONSTANT TCOLDCFA
11
12 0 CONSTANT ROOF      ( These constants are initialized by TARGET)
13 0 CONSTANT NLIST
14 0 CONSTANT TLIST
15 0 CONSTANT HEADERS
16 0 CONSTANT TREETAGCFA DECIMAL


SCR #124
 1 ( CONSTANTS for detection of special cases)
 2 ' (+LOOP) CFA CONSTANT (+LOOP)CFA
 3 ' LIT     CFA CONSTANT LITCFA
 4 ' (.")    CFA CONSTANT (.")CFA
 5 ' (ABORT") CFA CONSTANT (ABORT")CFA
 6 ' EXIT    CFA CONSTANT EXITCFA
 7 ' OBRANCH CFA CONSTANT OBRANCHCFA
 8 ' (LOOP)  CFA CONSTANT (LOOP)CFA
 9 ' (/LOOP) CFA CONSTANT (/LOOP)CFA
10 ' (;CODE) CFA CONSTANT (;CODE)CFA
11 ' BRANCH  CFA CONSTANT BRANCHCFA
12 ' :       CFA @ CONSTANT DOCOL
13 ' SP0     CFA @ CONSTANT DOUSE
14 ' 0       CFA @ CONSTANT DOCON
15 ' PREV    CFA @ CONSTANT DOVAR
16 ' FORTH   CFA @ CONSTANT DOVOC
```

```
SCR #125
 1 ( STRING PUT$ from A.Winfield "The Complete FORTH" Sigma 1983)
 2 HEX
 3 : STRING ( n --- )
 4 CREATE DUP C, 0 C, ALLOT
 5 DOES) 2+ DUP 1- C@ ;
 6
 7 : PUT$ ( a n --- )
 8 DROP 1-
 9 DUP 1- C@
10 24 WORD DROP
11 HERE C@ (
12 IF ." String too big" DROP QUIT THEN
13 HERE DUP C@ 1+
14 ROT SWAP CMOVE ;
15 DECIMAL
16

SCR #126
 1 ( TVRALLOT TVARIABLE TPICK TROLL ;S FLIP) HEX
 2
 3 : TVRALLOT ( n --- ) TVRAMOFF +! ;
 4
 5 : TVARIABLE ( --- ) TVRAMOFF DUP @ 2 ROT +! TVRAM + CONSTANT ;
 6
 7 : TPICK ( n1 --- n2 ) 2* SP@ + @ ;
 8
 9 : TROLL ( n --- ) 1+ DUP TPICK SWAP 2* SP@ +
10 BEGIN DUP 2- @ OVER ! 2- SP@ OVER U< NOT UNTIL DDROP ;
11
12 : ;S ( --- ) 0 'STREAM ! ; DECIMAL
13
14 ;S   HEX
15 : FLIP ( n --- n )
16 0 100 U/MOD SWAP 100 * + ; DECIMAL
```

```
SCR #127
 1 ( EAKER CASE STATEMENT FROM CEE 9/80 FORTH DIMENSIONS II/3 P 37)
 2
 3 : CASE   ?COMP CSP @ SP@ CSP ! 4 ; IMMEDIATE
 4
 5 : OF  4 ?PAIRS COMPILE OVER COMPILE = COMPILE OBRANCH HERE 0
 6      , COMPILE DROP 5 ; IMMEDIATE
 7
 8 : ENDOF  5 ?PAIRS COMPILE BRANCH HERE 0 , SWAP 2
 9      [COMPILE] THEN 4 ; IMMEDIATE
10
11 : ENDCASE  4 ?PAIRS COMPILE DROP BEGIN SP@ CSP @ = 0=
12      WHILE 2   [COMPILE] THEN REPEAT CSP ! ; IMMEDIATE
13
14 : ASCII BL WORD 1+ C@ [COMPILE] LITERAL ; IMMEDIATE
15
16

SCR #128
 1 ( F$ M$ FILE IMG TARG MAP DMA FCB LB) HEX
 2
 3 B STRING F$ F$ PUT$ TARGET  TMP$
 4 3 STRING M$ M$ PUT$ MAP$
 5
 6 : FILE ( --- ) CREATE HERE A2 ALLOT A2 0 FILL ;
 7
 8 FILE IMG
 9 FILE TARG
10 FILE MAP
11
12 : DMA ( fa --- a ) 1+ ;
13 : FCB ( fa --- a ) 81 + ;
14 : LB ( n --- b ) FF AND ;
15 DECIMAL
16
```

(Continued on next page)

Forth Compiler Listing (Listing continued, text begins on page 52)

```
SCR #129                                              12
  1 ( T'STREAM TWORD FILENAME?) HEX                  13 : FCREATE  ( fa --- f )
  2 : T'STREAM  ( --- a ) TIB @ )IN @ + ;            14 DUP 81 0 FILL FCB 16 SWAP SYSCALL LB ISFF ;
  3                                                   15 DECIMAL
  4 : TWORD  ( b --- a ) T'STREAM SWAP ENCLOSE DDUP ) 16
  5 IF DDROP DDROP 0 HERE !
  6 ELSE )IN +! OVER - DUP )R HERE C! + HERE 1+ R) 1+ CMOVE
  7 THEN HERE ;                                       SCR #131
  8                                                    1 ( FOPEN FCLOSE SETDMA FREAD FWRITE) HEX
  9 : FILENAME? ( fa --- )                             2
 10 ." ? " QUERY                                       3 : FOPEN ( fa --- f )
 11 TIB @ 1+ C@ ASCII : =          ( default or select?)  4   DUP 80 SWAP C! FCB F SWAP SYSCALL LB ISFF ;
 12 IF  3A TWORD 1+ C@ 40 - 0 MAX OF MIN OVER FCB C! ( select disk)  5
 13 ELSE 0 OVER FCB C! THEN         ( default disk)    6 : FCLOSE  ( fa --- f )
 14 DUP FCB 0C + 15 0 FILL DUP FCB 1+ 0B BLANK ( clear FCB)  7 FCB 10 SWAP SYSCALL LB ISFF ;
 15 2E TWORD COUNT 8 MIN 3 TPICK FCB 1+ SWAP CMOVE  ( load FCB)  8
 16 BL TWORD COUNT 3 MIN 3 TROLL FCB 9 + SWAP CMOVE ; DECIMAL  9 : SETDMA  ( fa --- )
                                                      10 DMA 1A SWAP SYSCALL DROP ;
                                                      11
SCR #130                                              12 : FREAD   ( fa --- f )
  1 ( ISFF FSEARCH FNSEARCH FDELETE FCREATE) HEX      13 FCB 14 SWAP SYSCALL LB ISZERO ;
  2                                                   14
  3 : ISFF  ( f --- f ) FF = IF FFFF ELSE 0 THEN ;    15 : FWRITE  ( fa --- f )
  4                                                    16 FCB 15 SWAP SYSCALL LB ISZERO ; DECIMAL
  5 : ISZERO  ( f --- f ) 0= IF 0 ELSE FFFF THEN ;
  6                                                   SCR #132
  7 : FSEARCH  ( fa --- f ) FCB 11 SWAP SYSCALL LB ISFF ;  1 ( SETDMA FREAD FWRITE FILE@ FILE! FILEEOF) HEX
  8                                                    2
  9 : FNSEARCH ( fa --- f ) FCB 12 SWAP SYSCALL LB ISFF ;  3 : FILE@  ( fa --- f )
 10                                                    4 DUP SETDMA DUP FREAD IF FFFF FF ELSE 0 0 THEN ROT C! ;
 11 : FDELETE  ( fa --- f ) FCB 13 SWAP SYSCALL LB ISFF ;
```

```
 5
 6 : FILE!  ( fa --- f )  DUP C@
 7   IF DUP SETDMA DUP FWRITE SWAP 81 0 FILL ELSE DROP 0 THEN ;
 8
 9 : FILEEOF ( fa --- f )
10   C@ FF = IF FFFF ELSE 0 THEN ;
11
12 : FILECREATE  ( fa --- f )
13   DUP FSEARCH NOT IF DUP FDELETE DROP THEN FCREATE ;
14 DECIMAL
15
16


SCR #133
 1 ( FC@ FC! IMG TARG MAP F$ M$) HEX
 2
 3 : FC@   ( fa --- b )
 4   DUP C@ 80 = IF DUP FILE@ DROP THEN
 5   DUP C@ 1+ DUP 3 TPICK C! + C@ ;
 6
 7 : FC!   ( b fa --- )
 8   DUP C@ 80 = IF DUP FILE! DROP THEN
 9   DUP C@ 1+ DUP 3 TPICK C! + C! ;
10 DECIMAL
11
12
13
14
15
16
```

```
SCR #134
 1 ( SPACEFILE SPACESFILE TYPEFILE U.RFILE) HEX
 2 : SPACEFILE  ( fa --- )  20 SWAP FC! ;
 3
 4 : SPACESFILE  ( n fa --- )  SWAP 0 MAX ?DUP
 5   IF 0 DO DUP SPACEFILE LOOP THEN DROP ;
 6
 7 : CRFILE  ( fa --- )  D OVER FC! A SWAP FC! ;
 8
 9 : TYPEFILE  ( a n fa --- )  ROT ROT DUP 0>
10   IF OVER + SWAP DO I C@ 7F AND OVER FC! 1 /LOOP
11   ELSE DDROP THEN DROP ;
12
13 : U.RFILE  ( # r fa --- )  ROT ROT
14   0 SWAP >R SWAP OVER DUP D+-
15   (# #S ROT SIGN #) R> OVER -
16   4 ROLL DUP ROT SWAP SPACESFILE TYPEFILE ; DECIMAL


SCR #135
 1 ( ID.FILE MAXNFA) HEX
 2
 3 : ID.FILE  ( nfa fa --- )  SWAP COUNT 1F AND ROT TYPEFILE ;
 4
 5 : MAXNFA   ( nfa1 a nfa2 --- nfa3 a ) [ SMUDGE ]
 6   BEGIN DUP DUP 0= SWAP @ A081 = OR NOT
 7   WHILE DUP PFA DUP CFA @ DOVOC = SWAP 4 + @ 0= NOT AND
 8     IF ROT ROT 3 PICK PFA 2+ @ MAXNFA ROT THEN
 9     OVER OVER U< NOT
10     IF ROT DDUP U< NOT IF DROP DUP THEN ROT ROT THEN
11     PFA 4 - @
```

(Continued on next page)

```
12  REPEAT DROP [ SMUDGE ] ;
13 DECIMAL
14
15
16


SCR #136
 1 ( NFTAG VERBTAG SFLAG )TTAG) HEX
 2
 3 : NFTAG ( nfa --- )
 4 DUP C@ 20 OR SWAP C! ;
 5
 6 : VERBTAG ( cfa --- )
 7 DUP 2- @ (;CODE)CFA =
 8 IF 0 SWAP [COMPILE] FORTH CONTEXT @ @ MAXNFA DROP
 9 ELSE 2+ NFA
10 THEN NFTAG ;
11
12 : SFLAG ( f --- f ) SWAP 1- SWAP ;
13
14 : )TTAG ( --- ) TREETAGCFA EXECUTE ;
15 DECIMAL
16


SCR #137
 1 ( PFSKIPS)
 2
 3 : PFSKIPS ( a1 --- a1 / a2 0 )
```

```
 4   DUP @
 5   CASE LITCFA OF LITCFA )TTAG 2+ 0 ENDOF
 6        (.")CFA OF (.")CFA )TTAG 2+ COUNT + 2- 0 ENDOF
 7        (ABORT")CFA OF (ABORT")CFA )TTAG 2+ COUNT + 2- 0 ENDOF
 8        OBRANCHCFA OF OBRANCHCFA )TTAG 2+ 0 ENDOF
 9        BRANCHCFA OF BRANCHCFA )TTAG 2+ 0 ENDOF
10        (LOOP)CFA OF (LOOP)CFA )TTAG 2+ 0 ENDOF
11        (/LOOP)CFA OF (/LOOP)CFA )TTAG 2+ 0 ENDOF
12        (+LOOP)CFA OF (+LOOP)CFA )TTAG (LOOP)CFA )TTAG 2+ 0 ENDOF
13   ENDCASE ;
14
15
16


SCR #138
 1 ( CREATE-DOES)) HEX
 2
 3 : CREATE-DOES) ( cfa --- )
 4   @ DUP VERBTAG DUP C@ DOESJMP =
 5   IF ' DOES) CFA VERBTAG DJMPLEN 2- +
 6      BEGIN 2+ DUP @ EXITCFA = NOT
 7      WHILE PFSKIPS ?DUP IF DUP @ )TTAG THEN
 8      REPEAT
 9   ELSE DUP DUP HEX CR U. 2+ NFA ID.
10      ." ---Code verb assumed." CR DECIMAL
11   THEN DROP ;
12 DECIMAL
13
```

```
14
15
16

SCR #139
 1 ( TREETAG) HEX
 2 : TREETAG ( cfa --- ) [ SMUDGE ]
 3 DUP 2+ NFA C@ 20 AND NOT
 4 IF 0 SWAP DUP 2+ DOCODE ! DUP VERBTAG DUP @
 5   CASE DOCODE @ OF SFLAG ENDOF
 6     DOCOL OF
 7       BEGIN 2+ DUP @ DUP EXITCFA = SWAP (;CODE)CFA = OR NOT
 8       WHILE PFSKIPS ?DUP IF DUP @ TREETAG THEN
 9       REPEAT SFLAG ENDOF
10     DOUSE OF DUP 2+ @ DUP 14 ) SWAP 36 ( AND
11       IF DUP EXECUTE @ TREETAG THEN SFLAG ENDOF
12     DOVAR OF SFLAG ENDOF
13     DOCON OF SFLAG ENDOF
14   ENDCASE SWAP NOT
15   IF DUP CREATE-DOES) THEN
16   THEN DROP [ SMUDGE ] ; DECIMAL


SCR #140
 1 ( TFIX;CODE X TFIXNUL APPTAG) HEX
 2 : TFIX;CODE ( --- )
 3 ' CREATE CFA VERBTAG ' CONSTANT CFA VERBTAG
 4 ' : CFA VERBTAG ' USER CFA VERBTAG ' EXIT CFA VERBTAG
 5 HEADERS IF ' BYE CFA TREETAG THEN ;
 6
 7 : X R) DROP ;
 8 : TFIXNULL ( --- )
```

```
 9 ' X CFA TREETAG 80E0 ' X NFA ! ;        ( patch for null)
10
11 : APPTAG ( --- ) CR CR ." Start tagging."
12 TFIX;CODE TCOLDCFA TREETAG
13 HEADERS ?DUP IF TREETAG TFIXNULL THEN
14 ."  Finished tagging." ; DECIMAL
15
16 ' TREETAG CFA ' TREETAGCFA !


SCR #141
 1 ( HITKEY DDRIVE LOADFCB IMAGESAVE) HEX
 2
 3 : HITKEY ( --- ) ."  and hit any key. --- " KEY DROP ;
 4
 5 : DDRIVE ( --- )
 6 CR CR ." Insert DOS disk in default drive," HITKEY ;
 7
 8 : LOADFCB ( a n fcba --- )
 9 DUP 21 0 FILL 1+ SWAP CMOVE ;
10
11 : IMAGESAVE ( --- ) DDRIVE
12 F$ IMG FCB LOADFCB
13 IMG FILECREATE DROP
14 HERE FORTHORG DO I 1- SETDMA IMG FWRITE DROP 80 /LOOP
15 IMG FCLOSE DROP CR CR ." Step 1 complete." DDRIVE BYE ;
16 DECIMAL


SCR #142
 1 ( FRAMEIN RAZE) HEX
 2
 3 : FRAMEIN ( --- )
```

(Continued on next page)

The Journal of Forth Application and Research

The aim of the *Journal* is to provide a reliable source of state of the art techniques and applications of Forth to scientific and industrial problems. The *Journal's* editorial review board is drawn from the foremost workers in Forth in the U.S., Europe and Canada. Past issues had in-depth coverage of such topics as Robotics, Data Structures, Forth Computers, Real-Time Systems, and Extended Address Computing. The *Journal* is open to all new work in Forth as well as the entire range of threaded interpretive languages and Forth-like systems.

The *Journal* is beginning its third year of publication. Published quarterly, the *Journal* contains applications and techniques papers, technical notes, review papers, algorithms, book reviews, and conference abstracts.

1985 Rochester Forth Conference Proceedings
The Proceedings of the 1985 Rochester Forth Conference will appear as a Special Issue of Volume 3 of the *Journal*. The conference had a theme of software management and engineering and how to improve software productivity. Invited papers discuss a space shuttle experiment using Forth, the automation of an airport with Forth, and Forth in the development of MAGIC/L. (Also available as a single issue for $20.; outside North America, $25.)

Subscriptions Volume 3, 1985 — Corporations and Institutions, $100.; Individual subscribers, $40. Please add $20 for airmail delivery outside N. America, and make checks in US funds from a US bank or international money order, payable to the *Journal of Forth Application and Research*, P.O. Box 27686, Rochester, NY 14627.

Forth Compiler Listing

```
4    SP@ 2- 2- ' ROOF !
5    VERBS# @ 2+ 0 DO 0 LOOP
6    SP@ ' NLIST !
7    HERE ROOF 2+ ! FFFF NLIST !
8    TVERBS# @ 3 + 2 * 0 DO 0 LOOP
9    SP@ ' TLIST !
10   FORTHORG DUP TLIST !
11   TARGORG - TLIST 2+ !
12   HERE NLIST 8 - !
13   FFFF NLIST 4 - ! ;
14
15 : RAZE  ( --- )
16   VERBS# @ 2+ TVERBS# @ 3 + 2 * + 0 DO DROP LOOP ; DECIMAL


SCR #143
 1 ( COUNTVERBS NFALISTFILL) HEX
 2
 3 : COUNTVERBS   ( nfa --- ) [ SMUDGE ]
 4   BEGIN DUP DUP 0= SWAP @ A081 = OR NOT
 5   WHILE DUP PFA DUP CFA @ DOVOC = SWAP 4 + @ 0= NOT AND
 6     IF DUP PFA 2+ @ COUNTVERBS THEN
 7     DUP C@ 20 AND IF 1 TVERBS# +! THEN PFA 4 - @ 1 VERBS# +!
 8   REPEAT DROP [ SMUDGE ] ;
 9
10 : NFALISTFILL   ( nfa --- ) [ SMUDGE ]
11   BEGIN DUP DUP 0= SWAP @ A081 = OR NOT
12   WHILE DUP PFA DUP CFA @ DOVOC = SWAP 4 + @ 0= NOT AND
13     IF DUP PFA 2+ @ NFALISTFILL THEN
```

```
14    DUP PFA 4 - @ SWAP DOWN @ ! -2 DOWN +!
15 REPEAT DROP [ SMUDGE ] ;
16 DECIMAL


SCR #144
 1 ( WINDOWTOP FRAMEBOTTOM) HEX
 2 : WINDOWTOP   ( a1 --- a2 )
 3   BEGIN DUP DUP @ SWAP 2- @ U< NOT
 4   WHILE 2-
 5   REPEAT 2- ;
 6
 7 : FRAMEBOTTOM   ( a1 n --- a2 ) SWAP
 8   BEGIN DUP @ 3 PICK U< NOT
 9   WHILE 2-
10   REPEAT SWAP DROP ;
11
12 : WINDOWBOTTOM   ( a1 a2 --- a3 ) @ OVER
13   BEGIN DUP @ DUP 4 PICK U< NOT SWAP FFFF = NOT AND
14   WHILE 2-
15   REPEAT SWAP DROP SWAP WINDOWTOP DDUP U<
16   IF SWAP THEN DROP ; DECIMAL


SCR #145
 1 ( OPENMOVE WINDOWSORT) HEX
 2
 3 : OPENMOVE   ( a1 a2 a3 --- )
 4   DUP DUP 2+ SWAP 5 PICK SWAP - DUP )R PREV @ SWAP BMOVE
 5   3 PICK 2+ SWAP 2+ 3 PICK 5 ROLL - BMOVE
```

```
 6  R@ - 2+ PREV @ SWAP R) BMOVE ;
 7
 8 : WINDOWSORT ( a --- )
 9  BEGIN DUP WINDOWTOP DUP @ FFFF = NOT
10  WHILE OVER OVER @ FRAMEBOTTOM
11     DDUP WINDOWBOTTOM OPENMOVE
12  REPEAT DROP DROP FORTHORG NLIST ! ;
13 DECIMAL
14
15
16


SCR #146
 1 ( FINDTAG OFFSETS) HEX
 2
 3 : FINDTAG ( a1 --- a2 )
 4  BEGIN 2+ DUP @ C@ 20 AND UNTIL ;
 5
 6 : OFFSETS ( --- ) CR CR ." Calculating offsets."
 7  NLIST 2+ @ TLIST NLIST TVERBS# @ 0
 8  DO FINDTAG DUP )R @ HEADERS NOT IF PFA CFA THEN
 9     ROT - SWAP 4 + SWAP OVER 2- @ + OVER
10     2+ ! R@ @ HEADERS NOT IF PFA CFA THEN OVER !
11     R@ 2+ @ SWAP R)
12  LOOP
13  DROP DUP 2+ @ SWAP 6 + ! DROP
14  ."  Offsets calculated." ;
15 DECIMAL
16


SCR #147
 1 ( VERBSLIST FIXSMUDGE) HEX
 2
```

```
 3 : VERBLISTS ( --- ) 0 VERBS# ! 0 TVERBS# !
 4  CR CR ." Creating verb lists."
 5  [COMPILE] FORTH CONTEXT @ @ COUNTVERBS FRAMEIN
 6  ROOF DOWN ! CONTEXT @ @ NFALISTFILL
 7  1 BUFFER DROP ."  Sorting." ROOF WINDOWSORT
 8  DECIMAL CR CR VERBS# @ U. ." (DEC) Verbs in host Forth."
 9  CR TVERBS# @ U. ." (DEC) Verbs in target code."
10  OFFSETS ;
11
12 : FIXSMUDGE ( --- )
13  TLIST 4 + TVERBS# @ 0
14  DO DUP I 4 + + @ HEADERS NOT IF 2+ NFA THEN
15     DUP C@ DF AND SWAP C!
16  LOOP DROP ; DECIMAL


SCR #148
 1 ( NEWLINK TARGMAP DUMP/=) HEX
 2
 3 : NEWLINK ( nfa1 --- nfa2 )
 4  TLIST 2- 2-
 5  BEGIN 2+ 2+ DDUP @ = UNTIL
 6  2- 2- @ SWAP DROP 0 FIXLINK? ! ;
 7
 8 : TARGMAP ( n nfa --- )
 9  SWAP TARGORG + DUP 4 HEX U.R SPACE MAPFILE
10  IF 4 MAP U.RFILE MAP SPACEFILE DUP MAP ID.FILE MAP CRFILE
11  ELSE DROP THEN OUT @ SWAP ID. OUT @ - A + DECIMAL
12  BEGIN 1- DUP 0< NOT WHILE SPACE REPEAT DROP
13  50 OUT @ F + < IF CR 0 OUT ! THEN ;
14
15 : DUMP/= ( a --- ) CR 20 - 50 DUMP TARG FILE! DROP
16  TARG FCLOSE DROP IMG FCLOSE DROP FIXSMUDGE ABORT ; DECIMAL
```

(Continued on next page)

```
SCR #149
 1 ( PATCHADD VAR) HEX
 2
 3 : PATCHADD  ( a1 --- a2 f ) DUP
 4 CASE TCOLD1   OF  TCOLDCFA 1 ENDOF
 5    INIT-USER  OF  TSP0     0 ENDOF
 6    U8         OF  TR0      0 ENDOF
 7    LTIB       OF  TSP0     0 ENDOF
 8    UP         OF  TR0      0 ENDOF
 9    RPP        OF  TR0      0 ENDOF
10    DUP        OF  DUP @    1 ENDOF
11 ENDCASE ROT DROP ;
12
13 : VAR ( a --- ) DOVAR? @ NOT OVER RPP = NOT AND
14 IF CR ." Images not matched at " DUP HEX 4 U.R DECIMAL DUMP/=
15 THEN PATCHADD DROP DUP FF AND TARG FC! FLIP FF AND TARG FC! ;
16 DECIMAL


SCR #150
 1 ( NEXTNFA REPAIRADD) HEX
 2
 3 : NEXTNFA  ( nfa1 --- nfa2 )
 4 NLIST 2- BEGIN 2+ DDUP @ U< UNTIL @  SWAP DROP ;
 5
 6 : REPAIRADD  ( a1 --- a2 )
 7 DUP @ DOVAR? = IF 2 DOVAR? ! THEN
 8 DUP PATCHMAX < IF PATCHADD ELSE @ 1 THEN
 9 IF FIXLINK? @ ?DUP IF SWAP DROP NEWLINK THEN
10 TLIST
```

```
11    BEGIN DDUP @ U< NOT
12    WHILE 2+ 2+
13      REPEAT 2- @ -
14 THEN ;
15 DECIMAL
16


SCR #151
 1 ( VERB)TARG) HEX
 2
 3 : VERB)TARG  ( nfa --- )
 4 DUP NEXTNFA SWAP
 5 DO IMG FC@ DUP I C@ =
 6    IF I FIXLINK? @ =
 7      IF DF AND !FIX @ IF 0 FIXLINK? ! 0 !FIX ! THEN
 8        THEN TARG FC! 1
 9    ELSE IMG FC@ 100 * + 201 - I @ =
10      IF I REPAIRADD DUP FF AND TARG FC!
11         FLIP FF AND TARG FC! 2
12      ELSE I VAR 2
13      THEN
14    THEN DOVAR? @ IF -1 DOVAR? +! THEN
15 /LOOP ;
16 DECIMAL


SCR #152
 1 ( DISCARD RELOCATE) HEX
 2
 3 : DISCARD  ( n --- )
```

```
4    ?DUP IF IMG SWAP 0 DO DUP FC@ DROP LOOP DROP THEN ;
5
6  : RELOCATE  ( --- n )  0 DOVAR? !
7    0 OUT ! 0 !FIX ! 0 FIXLINK? !
8    CR CR TARGORG 4 HEX U.R ." HEX (target origin)" CR CR
9    MAPFILE IF TARGORG 4 MAP U.RFILE MAP CRFILE THEN
10   FORTHORG 0 NLIST 8 - TLIST
11   DO I TLIST = NOT IF DUP I @ TARGMAP I @ FIXLINK? ! THEN
12      I @ [ ' ! NFA ] LITERAL = IF 1 !FIX ! THEN
13      SWAP I @ SWAP - DISCARD
14      I @ VERB)TARG I @ DUP NEXTNFA DUP ROT - ROT + 4
15   /LOOP SWAP DROP DECIMAL ;
16 DECIMAL
```

```
SCR #153
1 ( VERB)TARGNH) HEX
2
3 : VERB)TARGNH  ( cfa --- )
4   DUP NEXTNFA SWAP
5   DO IMG FC@ DUP 1 C@ =
6     IF TARG FC! 1
7     ELSE IMG FC@ 100 * + 201 - I @ =
8       IF I REPAIRADD DUP FF AND TARG FC!
9         FLIP FF AND TARG FC! 2
10      ELSE I VAR 2
11      THEN
12    THEN DOVAR? @ IF -1 DOVAR? +! THEN
13  /LOOP ;
14 DECIMAL
15
16
```

```
SCR #154
1 ( RELOCATENH) HEX
2
3 : RELOCATENH  ( --- n )
4   0 OUT ! 0 FIXLINK? ! 0 DOVAR? !
5   CR CR TARGORG 4 HEX U.R ." HEX (target origin)" CR CR
6   MAPFILE IF TARGORG 4 MAP U.RFILE MAP CRFILE THEN
7   FORTHORG 0 NLIST 8 - TLIST
8   DO I TLIST = NOT IF DUP I @ 2+ NFA TARGMAP THEN
9     SWAP I @ SWAP - DISCARD
10    I @ VERB)TARGNH
11    I @ DUP NEXTNFA
12    DUP ROT - ROT + 4
13  /LOOP SWAP DROP DECIMAL ;
14 DECIMAL
15
16
```

```
SCR #155
1 ( SETUPDOS CLEANUPDOS) HEX
2
3 : SETUPDOS  ( --- )
4   DDRIVE F$ IMG FCB LOADFCB
5   BEGIN IMG FSEARCH
6   WHILE CR ." TARGET.TMP not found on default drive" DDRIVE
7   REPEAT IMG FOPEN DROP
8   CR CR ." Target code filename"
9   TARG FILENAME? TARG FILECREATE DROP MAPFILE
10  IF TARG FCB MAP FCB 21 CMOVE
11     M$ MAP FCB 9 + SWAP CMOVE MAP FILECREATE DROP
12  THEN ;
```

```
13
14 : CLEANUPDOS  ( --- ) IMG FCLOSE DROP
15 IMG FDELETE DROP TARG FILE! DROP TARG FCLOSE DROP MAPFILE
16 IF 1A MAP FC! MAP FILE! DROP MAP FCLOSE DROP THEN ; DECIMAL


SCR #156
1 ( TARGETMEMORY)
2
3 : TARGETMEMORY  ( n --- ) CR CR
4 ." (HEX) Target Memory Map" HEX
5 CR TVRAM TVRAMOFF @ + 1- 4 U.R ."  Last needed byte of RAM."
6 CR TVRAM 4 U.R ."  Target variable RAM start {TVRAM}."
7 CR TEM 4 U.R ."  Target-end-of-memory {TEM}."
8 CR TBUF1 4 U.R ."  Start of block buffer area {TBUF1}."
9 CR TR0 4 U.R ."  Initial return stack pointer {TR0}."
10 CR TSP0 4 U.R ."  Initial stack pointer {TSP0}."
11 CR TARGORG + 1- 4 U.R ."  Target code end address {last used}."
12 CR TARGORG 4 U.R ."  Target origin {TARGORG}."
13 DECIMAL ;
14
15
16


SCR #157
1 ( TARGET) HEX
2
```

```
3 : TARGET  ( pfa or 0 --- )
4 DUP IF CFA THEN ' HEADERS !
5 TCOLDCFA 2+ NFA FENCE ! FREEZE
6 APPTAG
7 FORTHORG 301 = IF IMAGESAVE THEN
8 VERBLISTS
9 SETUPDOS
10 HEADERS IF RELOCATE ELSE RELOCATENH THEN
11 CR CR HEX DUP U. DECIMAL ."  (HEX) Bytes in target. "
12 TARGETMEMORY CLEANUPDOS FIXSMUDGE RAZE
13 CR CR ." Target compilation complete." ;
14 DECIMAL
15
16
```

End Listing

CRAFTSMAN™ QUALITY SOFTWARE FOR 'C' & OTHER PROGRAMMERS

C-TREE
B-Tree File Manager, Source Code, No Royalties!

NEW FEATURES!

A b-tree can be infested with bugs, so before buying one, ask its age. In a stand of saplings, this one is a real cequoia. C-tree™ has been around since 1979. (It became Digital Research's Access Manager™). That means seasoned, sturdy code which hasn't cracked under prolonged and widespread use.

C-tree comes in C source code, revealing all you've ever wanted to know about how b-trees are written. Provided you bind it into your binary application, you can re-distribute c-tree without royalties.

And if all this is disappointing, now the good part. C-tree's design splits nodes to allow any number of users to access an index file simultaneously even when updates are in progress so that multi-user configurations and adaptation to networks are possible.

The latest version has new features: support of variable record length data files; multiple key indexes in a single physical file; MS-DOS and Unix record locking examples.

Thanks to source code which does not deviate from K&R, C-tree can travel. Binary has always meant finding a substitute file manager for yet another compiler, operating system, or computer; then changing all the function calls program-wide; then starting the whole testing process anew. That's over. Tests in many environments prove that C-tree gives your application a ticket to anywhere.

C-tree permits any number of keys for a data file, supports duplicate keys, alphanumeric or numeric, etc., etc.: it's a big product with everything you'd expect. Beyond that it is intelligently designed as both a high level set of ISAM routines to minimize your coding in handling all details of adding a record on its own, for example; and as low level operations which you can access directly. Either way C-tree maintains optimal index structures which will find a record amongst a million ten byte keys in no more than five disk seeks.

Product Code: F0660(List Price: $395 Our Price: $329

TEXT TOOLBOX
Tackles Text Tangles

Unix™ boasts a number of muscular utilities that are migrating to the PC world. Lattice has assembled a cluster of the most useful text management tools into a single package.

"Grep" looks for text patterns in any number of files. Want all occurrences of a global variable throughout a program system? Want to search all programs in a directory, down paths to other directories, or all files on a disk? Need to find all the function calls in an entire program system? Grep can do it with a powerful expression syntax that goes far beyond your text editor's search command.

"Ed" is similar to the well-known Unix editor. It offers search and replace with "grep's" syntax, block move, read and write, optional line numbering, append, insert, delete, and this unusual facility: you can instruct "ed" to apply a file of commands to any number of target files, even complicated changes and text additions, such as those created by "diff".

"Diff" compares text files line for line. Its output is a precise list of instructions telling what to do to make two files the same, a list which can be handed to "ed" to do.

Code:	List Price:	Our Price:
L2200	$120	$100
L2205 with Source	$240	$200

CVUE
Make Your Own Editor

CVUE is a low-priced screen-oriented text editor which does most of the things that a good editor should do, and boasts full DOS 2.0 directory path name support in reading and writing files.

It was written by the Lattice programmers who felt forgotten by the folks who write WP software. They needed easy entry of non-display characters such as control codes and escape sequences, not footnotes; indenting and undenting, not italics; pattern searching, not spell checking. So CVUE™ was born.

CVUE only supports in-memory text files, but with memory at today's prices, creating and maintaining files of over 500 K is practical. As compensation, CVUE is very compact and fast. It actually runs in computers with only 64 K of memory and uses no tediously slow overlays.

The power of CVUE is its ease of customization. And when you take advantage of the Source Code option, the resultant editor can be made truly your own.

Code:	List Price:	Our Price:
L2240	$75	$69
L2245 with Source	$200	$220

USED COMPILERS WELCOME
Trade In for the Latest Model Lattice C

Has your compiler run out of gas? Has your model been discontinued? Is it falling behind for lack of new parts?

Even if yours is in good shape, you have surely noticed there are more options and accessories produced to run with the Lattice C compiler than any other. Don't do without these additives any longer. It's time for new license plates. Trade in your original disks and manual of any of the compilers below and we'll send you Lattice's most up-to-date model.

From then on you will be adopted by Lattice for full, direct support by their technical specialists.

	Price:
Microsoft MS-DOS/PC-DOS C	$150
Computer Innovations C86, Mark Williams C, Digital Research C, Whitesmith's C	$200

PRE-C
Thorough "Lint"-like Analysis Now on the PC

Unix users long for a "lint" to give programs a thorough cleaning before they disappear into a compiler.

Pre-C™ looms larger than "lint." It finds problems your compiler won't. Problems that a debugger will have trouble figuring out. Even problems which will cause trouble with other compilers.

Pre-C finds all the syntactical tripwires that will blow out a compiler, sure, but it goes after subtler problems: code which will never be accessed, casts with suspect conversions, variables declared as external but never used, functions never called, obsolete usage (even C has changed), machine-dependent expressions with will inhibit portability.

Compilers work with one module at a time. They know nothing of other modules which only meet up at link time. Pre-C can look at all segments of your program at once and report to you any inconsistencies of inter-module references: conflicting data type declarations, parameter lists in function calls which disagree with the functions themselves in number or data type, declarations of external functions which differ from their definition.

Pre-C uses the Unix System III compiler standard to safeguard maxium portability anywhere in the C world. There are then plentiful command line options to advise Pre-C what to flag and what to forgive, useful during early coding when some functions are empty or incomplete. The output of each analysis can be filed for use with subsequent Pre-C runs, so work is not performed redundantly.

Pre-C lets you develop standing profiles of binary libraries. In any C program you subsequently write, Pre-C can use these profiles to make sure your calls to those libraries' functions are perfect.

This is a big product which will work miracles in speeding large system development. 128K minimum; 192K recommended.

Product Code: P0590	
List Price: $395	Our Price: $329

dBC
Switch from dBASE Language to C for Power, Speed

There are a lot of dBASE™ file users out there. Most of them just keep data bases and use dBASE's limited reporting facilities. They're not programmers, so they don't use the dBASE programming language. But they'd like more for their efforts, and that's a business opportunity.

dBC™ links C to dBASE. It creates and maintains files and their indexes which exactly replicate dBASE file design. So dBASE can read and update them. And the reverse. dBC can use any files created by dBASE. Now C and dBASE can operate on the same data bases interchangeably.

That opens up the widespread culture of dBASE installations to exploitation by C programmers. Now you can replace the resident dBASE language with the speed of C. And you no longer have to write every line of code, as in the dBASE language, because now you've unlocked C's vast storehouse of off-the-shelf libraries and utilities.

dBC's functions parallel all dBASE's file handling commands, many of them decomposed to give you closer-in control. The manual discusses each in detail, and demonstration source files on your disk show how every function is used. Use dBC for custom work for clients, or design generalized programs for manipulation and reporting of dBASE data bases.

Or use dBC on its own. It's a complete ISAM file manager for use with the Lattice C compiler whether or not dBASE will ever be used in tandem, has versions for all four memory models, and can have sixteen index and data files open at once.

Code & Version:	List Price:	Our Price:
L00II: Dbase II Compatible	$250	$225
LCCII: with Source Code	$500	$450
L0III: Dbase III Compatible	$250	$225
LCIII: with Source Code	$500	$450

Neon, Version 1.0

Company: Kriya Systems, Inc.
505 N. Lake Shore Drive,
Suite 5510, Chicago, IL
60611, (312) 822-0624
Computer: Apple Macintosh
(128K)
Price: $155.00
Circle Reader Service No. 101
Reviewed by Bruce Horn

This review is reprinted with the permission of the author, Bruce Horn, and The Club Mac News, *in which it originally appeared (Vol. 1, No. 12, May 1985).*

I have not written many programs in Forth. I've always considered Forth to be a sort of high-level assembly language, fun for hacking around in, but not suited to development of large programs. I have written Smalltalk programs, though, and when I heard that Charles Duff and the folks at Kriya had developed Neon, an object-oriented programming environment that was said to be a cross between Smalltalk and Forth, I was intrigued; is this the long-awaited favorite programming environment for the Macintosh?

History of Object-Oriented Programming

Object-oriented programming began in the early 1960's with the development of Simula by the Norwegian Computation Center in Oslo, Norway. Simula was the first language to implement the Class construct. In the early 1970's, the Learning Research Group at Xerox Palo Alto Research Center began implementation of the Smalltalk programming environment. Smalltalk was the first system to be designed completely around the Class/Object concept. Many other languages have since provided sup-

port for classes, objects, and subclassing, including CLU, Ada, C++ (an extended version of C), and several versions of LISP, though objects and classes are not as well-integrated into these languages as they are in Smalltalk. The ease of experimentation with objects was one of the key reasons that much of the interesting user interface designs borrowed for Macintosh and Lisa were created in the Smalltalk system. (See *Smalltalk-80: The Language and its Implementation,* by Adele Goldberg and Dave Robson, for more details.) The latest version of Smalltalk has been licensed to several universities and computer companies; Tektronix markets a 68000-based Smalltalk machine for around $14,000, and a version of Smalltalk-80 will soon be available for the Macintosh XL.

Classes and Objects—a Quick Overview

In Neon, classes define objects, a convenient packaging of 1) the data that defines the object and 2) the object's actions. For example, Class **Pen** defines what a pen object looks like (as far as its internal state) and acts like (the messages it responds to). The internal state, or instance variables of a **Pen** object, might include location, direction, and pen size; the messages that pens respond to would include **move:, turn:, goto:,** and **home:.** (The Neon tutorial has an excellent discussion of the implementation of a **Pen** class, with examples and line-by-line explanation of the source code). Class definitions are bracketed by **:Class** and **;Class** words, and include a list of instance variables (which may themselves be objects) and methods (the procedures run when a message is sent to the object).

Defining a class is like designing a little computer, with its own memory

(the instance variables) and instruction set (the messages it can receive). In using a computer, usually you don't concern yourself with how the computer executes the instructions you give it. If you did, and you counted on the various nuances of the implementation of each instruction, programming the computer would be much more complicated. Classes protect you from the need to know about the implementation of the object (computer) by allowing you to talk to the object through a set of messages.

Responses to messages include returning a result on the stack, changing the internal state of the object, sending messages to other objects, or just producing a side-effect (such as drawing a line on the screen). **10 move: Bic** sends the message **move:** to the pen object named **Bic,** changing its location instance variable and possibly drawing a line on the screen; **45 turn: Bic** turns the pen 45 degrees to the right. One advantage of having a pen *object* do your line drawing on the screen instead of just calling Quickdraw directly is that you can have many separate pens, each with its own state, and you can send messages to them independently. To paraphrase Dan Ingalls, one of the developers of Smalltalk, programming with objects is like working with trained animals instead of pushing data around with a broom.

Classes may be defined that are subclasses of another class. This causes objects of that class to inherit both the instance variables and actions of the superclass. Subclassing provides the capability of saying, "I want something just like that, except a little different." For example, you might want to have a **Pen** that recorded its movements. Since the current class **Pen** handles everything except the recording, subclassing **Pen**

and creating a **RecordingPen** that provides the capability would make the most sense.

Since class **RecordingPen** is a subclass of **Pen**, **RecordingPen** objects will have both **Pen** and **RecordingPen** instance variables, and will respond to the union of all **Pen** and **RecordingPen** messages. In general, if the message is not defined in the class that the object belongs to, Neon will search up the superclass chain to find a class that can respond to the message, and runs the appropriate method. You may also add new messages and instance variables to distinguish your class from the superclass, and you can override superclass messages by providing them in your subclass.

To create the new **RecordingPen** class all you would need to do is add **startRecording:**, **stopRecording:**, and **playBack:** messages, and a new instance variable **penRecord**, which keeps the list of messages sent to the object since recording started. By overriding each message that the object wants to record, it can remember the message in **penRecord** and send the appropriate message to its superclass, **Pen**, to do the actual drawing.

Building a system from scratch based on classes and objects provides an extremely flexible, compact and modifiable system. Since the code that defines the system itself is accessible, extensions to the system may be made quickly and easily at even the lowest level.

(For an excellent discussion of classes, messages and object oriented programming, see Brad Cox's article "Message/Object Programming: An Evolutionary Change in Programming Technology" in the January 1984 issue of *IEEE Software*.)

Built-in Classes

Since Neon, like Smalltalk, was written in itself, using its own object and class concepts, the groundwork for building applications is already there and fully accessible to the programmer. The basic classes provided in Neon make it easy to support the Macintosh User Interface standard with minimal effort. In Pascal, a Mac program typically has several large case statements at the very highest

level of the main program to classify events received and to dispatch them to handlers. Neon helps to take care of this by encapsulating most of the standard event handling in class **Event**, menu selection and dispatching in class **Menu**, and standard window actions in class (you guessed it) **Window**. Through these classes, Neon shields the programmer from the standard necessities and deals with most of the housekeeping that is required in a Macintosh application behind the scenes.

There are many other classes written to simplify interaction with the User Interface Toolbox. Class **File** provides access to the Toolbox File I/O and Save/Open dialogs, and class **String** defines messages for printing, insertion, and pattern matching of text strings (a nice interface to Munger, the Toolbox string handler). In addition, there is a set of nine classes to support Quickdraw, including **Rect**, **Oval**, **Picture**, and **GrafPort**. All Toolbox routines are directly accessible.

Using Neon

Neon is started by double-clicking on **neon.com**. **Neon.com** is a prebuilt environment in which a basic set of classes, messages, and objects have been installed. As you work in this environment, you will be creating classes, objects and messages, as well as new Neon words. Once you have modified the default environment you may save it on the disk. In this way you can build your own custom Neon system.

When Neon starts up, it comes up with a single teletype window, where you can type Neon commands. Output from some of the utilities appears there also. The Neon window does not update when windows are moved around above it, which is a minor nuisance. There is no text editing allowed in the Neon window other than backspacing.

It is fairly easy to create new windows and menus while in Neon and to attach Neon words to each menu item. For example, sending the message **Example:** to a new window creates a Neon window that will interpret any text typed to it. There are

also example messages in several other classes, including **Control**.

The standard **Neon.com** system, as it comes from Kriya, has five menus:

1) The **Apple** menu has the standard list of desk accessories, an "About Neon . . ." item, and an extra accessory called "Editor." This accessory provides the text-editing facilities necessary for editing class and object definitions within Neon.

2) The **File** menu consists of five items:

Load... loads a text source file into the current environment, defining new classes, objects, and messages. You would load **grdemo.load** to define the special classes for the graphics demonstration program.

Save saves your current environment onto the disk. As you modify classes and customize the system to meet your needs, you can save the complete system and return to it by simply opening the appropriate Neon document. This is often called "taking a snapshot."

Save As... allows you to save your current environment under a new name.

Print... prints a specified text file to the printer. The print command produces a formatted ("prettyprinted") listing of the source code in the file.

Quit closes all files and returns to the Finder. All changes made to the Neon environment are lost (unless they had been previously Saved.)

3) The **Edit** menu provides the standard **Undo**, **Cut**, **Paste**, **Copy**, and **Clear** items, as well as a special **Editor** item to start up the editor desk accessory.

4) The **Utilities** menu has six items:

List Words lists all words currently defined in the dictionary.

List Objects lists all objects of a given type that currently exist in the environment.

List Classes lists all defined classes in the current environment, indenting to show subclass relationships.

Examine Memory is a feature that allows you to examine and search the Macintosh memory.

Grep... searches all text files on the disk for a given string. The results of the search (file name, line number, and text line) are printed in the Neon window.

Install lets you change the sizes of the stack, dictionary, and heap. The stack and dictionary are adjustable, while the heap gets what is left.

5) Finally, the **Neon** menu provides five items:

Echo to Printer echoes all text that appears in the Neon window to the printer, including text that is typed by the programmer. This is especially useful in that it allows the programmer to keep a record of the input and output of a program being debugged.

Echo During Load displays the text from a source file being loaded in the Neon window as it is being compiled. This is useful in helping the programmer debug a new module or class definition.

Show Free Space displays the amount of memory available in the dictionary and the heap in the Neon window.

The last two entries in the Neon menu deal with modules. Modules are separately compilable, relocatable code segments that are loaded and purged from memory dynamically. This allows large applications to run without having to be completely resident in memory. These modules act similarly to the Macintosh code segments in that they are dynamically loaded into memory and purged when not in use. Many of the menu items in Neon are defined by modules, including **Grep** and **Examine**.

Show Module Status lists all purgeable modules currently defined in the dictionary and shows whether they are loaded by printing their memory locations (or 0 if not loaded).

Purge Modules purges all modules from the heap, freeing up space for a memory-intensive operation.

Because the Neon window provides no editing capability, the system comes with an editor to edit class definitions within Neon. This editor is an enhanced MockWrite desk accessory, from CE Software. I found it easy and convenient to use, though I would prefer a built-in editor in

which I could select, edit, and execute Neon statements directly. Once a class definition is created, it may be loaded into the system by either typing // followed by the filename or by selecting the **Load** item in the file menu. Either action begins compilation of the file. After the file has been completely compiled and defined in the system, you may create new objects, send messages to previously-created objects, and so on.

After a few hours with Neon, I felt quite at home creating objects, sending messages, and building graphical tools. The fact that Neon is an interactive programming environment is a great advantage in building applications for the Macintosh because the turn-around time is much shorter and experimentation is easier. Neon also provides a facility for "sealing off" an application, creating a double-clicking program without access to the Neon environment.

Forth Improvements

Besides support for objects and classes, Neon has other improvements to standard Forth. The main objection to Forth is usually the need for strange and unintelligible stack manipulation in function definitions; Neon avoids this through local variables and named parameters. Consider the example in the Neon tutorial for calculating $a^2 + b^2$:

```
:formula1 { a b -- solution }
a a * → a
b b *
a + . CR ;
```

The input parameters are named between the braces, and are pulled off the stack and stored in the local variables **a** and **b**. The arrow ($\rightarrow$) stores the value on the stack into the named parameter. Local variables may also be defined within the braces by preceding the variable list with a backslash. Again, from the Neon tutorial, consider the formula $(a + b - 3c)/(b + 3c)$:

```
:formula2 { a b c \ num den -- result }
```

```
a b + 3 c * → num
3 c * b + → den
num den / ;
```

The numerator and denominator are calculated separately, stored in local variables, and then the formula is calculated. With local variables and named parameters, Neon programs are downright readable.

Summary

Neon is nicely designed. It is clear that the developers had an excellent understanding of what object-oriented programming is supposed to be. With the basic classes, examples, and tutorial, you could be writing interesting programs that support the Macintosh User Interface Standard in a matter of weeks. This compares very favorably with the months that most people take to build the support routines that a real application requires. Kriya is also planning to support a clearinghouse service for useful, public domain classes written in Neon for use by software developers.

Neon is also a good programming language for occasional use, though it has so many features and is so flexible that it could take a little while to learn what is provided in the system. The documentation is thorough and well-written, and the tutorial is interesting and useful.

As a programming environment, Neon is very well suited to the Macintosh. The built-in classes are well designed and useful, and provide a much improved interface to the Macintosh Toolbox routines. It is flexible, powerful, fast and a lot of fun. It is also the only implementation of an object-oriented system for the Macintosh so far. Because of the support for classes and objects, organization of large programs would be quite manageable, while with standard Forth working on a large system with thousands of words might become confusing. The only disadvantage that Neon has, in my opinion, is the backwards Forth syntax. If somebody comes up with a new infix parser-scanner to allow Smalltalk-like expressions, Neon could be a real winner in the programming language sweepstakes.

MacFORTH Level 1, Version 1.2; Level 2, Version 2.1; Level 3, Version 3.0

**Company: Creative Solutions, Inc.
4701 Randolph Rd, Ste. 12
Rockville, MD 20852
(301) 984-0262**
Computer: Apple Macintosh (128K)
Price: Level 1—$149, Level 2—$249 (includes Level 1), Level 3—$499 (includes Levels 1 and 2)
(Circle Reader Service No. 103)
Reviewed by Alan Clute

The Macintosh is a giant toolkit. An ideal language to take advantage of this toolkit is Forth. MacFORTH, from Creative Solutions, Inc., is an excellent Forth for the Macintosh. MacFORTH delivers the richness and power of the Macintosh without getting overly complex. The package, while being simple to use, still leaves the programmer enough breathing room to develop programs that are powerful and make good use of the Mac's capabilities.

MacFORTH is a production language. Current products written in MacFORTH include PeachTree's *Back to Basics Accounting*, Brainworks' *ChipWits*, and Ideaform's *MacLabeler*. An interesting new product is CSI's *StudioMac*, a music program that includes tutorials, source code, MIDI interface capability, and complete plans to build an inexpensive MIDI interface for the Macintosh.

Contents of Package

Table 1 (see page 101) shows the main contents of each level of Mac-FORTH. The three levels are nested, and upgrades are available for the difference in list price. Because Level 2 includes more and is in some ways easier to work with than Level 1, it is probably the best package to get first, especially at mail-order prices. For the record, this review concerns itself with Versions 1.2, 2.1, kernel 2.3 and, briefly, with Level 3 version 3.0.

MacFORTH has been around a fairly long time, as Macintosh software goes. Level 1 was first released in

April of 1984, with the first version of Level 2 coming out in July of that year. MacFORTH was the first serious native language for the Macintosh. It makes good use of a 128K Macintosh and can be used with only one disk drive (if you are so inclined).

MacFORTH is not copy protected, so it can easily be backed up or installed on a hard disk. The accompanying license agreement says in no uncertain terms that you should not give or sell copies to others, but CSI is to be congratulated for not otherwise interfering with your use of the program.

Forth Implementation

MacFORTH is derived from the Forth-79 standard, optimized for the Macintosh environment. It is a 32-bit Forth, which is more suitable for the large address space of the Mac than 16-bit implementations. Considered just as Forth, MacFORTH is a good one. Converting programs from another standard version of Forth should be rather straightforward.

The internal implementation is direct threaded with a token table to keep all word references to 16 bits. Thus, program code requires about the same amount of dictionary space as in a 16-bit version of Forth. Variables occupy 32 bits. Fetch and store operators come in several flavors to support 8, 16, or 32 bit values.

MacFORTH uses traditional Forth blocks for source code, but these blocks are within block files. MacFORTH thus lives very comfortably within the Macintosh environment.

Programs can do disk I/O using block files, or can choose to use data files with text, fixed-length record, or byte-addressable virtual formats. The entire storage-management vocabulary includes about 90 words.

Documentation

MacFORTH is extremely well documented. Most of the documentation covers the Macintosh interface, but then, most of the Forth is quite standard. The level 1 manual is about 160 pages, plus glossary and appendices. Level 2 runs over 200 pages, plus glossary, and Level 3 adds 200 more pages, which show you step by step how to turn a running program into a production disk.

The manuals make clear presentations. The glossaries are well organized, well indexed, and well written. The only thing obviously missing is a reference in each glossary entry to the page where the word is discussed in context.

MacFORTH pretty much assumes you know Forth, although it comes with a Forth tutorial and a rich set of examples that would prove useful even to a beginner. As is often advised, the beginner really should arm him/herself with some instructional book like *Starting FORTH* or *Thinking FORTH* by Leo Brodie (Prentice-Hall, 1981 and 1984) instead of relying on a language manual. In addition, a new book should be out by the time you read this, *MacFORTH: Programming for the Rest of Us* by Dave Colburn, one of the authors of MacFORTH (also available from CSI).

The Going Forth tutorial, a split-screen affair with paged text on one side and the live MacFORTH system on the other, provides a hands-on introduction to MacFORTH. Source is on disk for those who want to expand the tutorial for others.

You get plenty of source code with these packages, both "real-world" stuff, like the block editor, and "tutorial" material like the Note Card

Level	Feature	Comments
1	• Going FORTH tutorial	Disk based, interactive.
	• Block-oriented editor	Similar to MacWrite.
	• Macintosh interface	Menus, windows, mouse, sound, event handler.
	• File system	Text, record, virtual, and block files. MacWrite text files.
	• Printer/Serial interface	Window or screen output, full serial I/O.
	• Programming aids	Trace, debug, error traps, direct calls to toolbox routines.
	• Source code	Forth extensions, editor, many examples, terminal program, download conversion, standard file-open dialogue.
2	• Level 1 included	
	• System tools	Heap management, 68000 assembler, snapshot (vocabulary save).
	• Advanced graphics	Polygons, regions, pictures, rectangle arithmetic and scrolling.
	• Text editing	TextRecords and TextFields.
	• Controls	Buttons, boxes, and scroll bars.
	• Clipboard I/O	Read and write from/to clipboard
	• Floating Point	IEEE-754 standard
	• Source code	Graphics demo, note card data management demo (fully annotated in manual), many examples, source for above extensions.
3	• Level 2 included	
	• Application workshop	Turnkey production tools and tutorial. Supports 512K Mac, overlays on 128K Mac.
	• Resource workshop	Macintosh resource customization and tutorial.
	• Kernel workshop	Forth kernel customization and tutorial.

Table 1:
Contents of MacFORTH Levels

example. This last is part of Level 2 and includes 23 screens of source and 20 pages of detailed manual text. In all, Level 1 includes about 120 screens of source, and level 2 about 140 screens.

Forth and the Macintosh

Forth has been described as a language for the development of problem-oriented languages. MacFORTH is a Macintosh-oriented language. The "problems" of the Macintosh—how to use windows, pull-down menus, graphics, and other features—have been smoothly addressed by MacFORTH.

The MacFORTH vocabulary is huge, over 900 words in Level 1 with another 250 or so added in Level 2. Most of these are the "application words" that give you direct control over the Macintosh. The Macintosh and its fabled toolbox are made directly available, letting you do things "the Macintosh way" with relative ease. Don't be intimidated by the large vocabulary: flexibility comes from lots of little words, rather than a few large ones, and the Macintosh has many of these small units. The window-handling package alone has over 70 words associated with it.

```
            ( create window )
15 CONSTANT testWindow
NEW.WINDOW testWindow

       ( define title )
" Example Window"                    testWindow W.TITLE
       ( define rectangle )
50 50 100 300                        testWindow W.BOUNDS
       ( add boxes )
CLOST.BOX SIZE.BOX +                 testWindow W.ATTRIBUTES
       ( add window to the screen )
testWindow ADD.WINDOW
```

Figure 1
Code to Create Window

```
           ( create menu, define title )
11 CONSTANT testMenu
0 " Example" testMenu NEW.MENU
         ( define item text )
" First;Second;Third(;" testMenu APPEND.ITEMS
         ( add menu to the screen )
DRAW.MENU.BAR
```

Figure 2
Menu Definition

```
        ( define menu handler )
testMenu MENU.SELECTION:
         ( CASE statement for items )
    CASE 1 OF FirstProgram ENDOF
         2 OF SecondProgram ENDOF
         3 OF ThirdProgram ENDOF
    ENDCASE
         ( turn off title hilite )
    0 HILITE.MENU ;
```

Figure 3
Selection of Menu Item

Examples

Straightforward use of windows and pull-down menus, for example, turns out to be simplicity itself. The code in Figure 1 (at left) creates a window with a close box and a size box. Figure 2 (at left) shows an example menu definition, with 3 items. The first two are "active" and the third is "inactive" (grayed out and incapable of being selected). The trailing (in **Third(** causes that item to be inactive. Type style (bold, underline, etc.) and command key equivalents are easily assigned using other character suffixes. Selection of a menu item is easily linked to the proper actions using a defining word (Figure 3, at left). Fancier uses of scrollbars, control buttons, and the like are implemented with the clarity and simplicity so essential to good Forth style.

The Level 2 manual includes a reference section giving the Pascal name, MacFORTH name, and stack usage for the over 150 toolbox traps directly callable by the user. For advanced applications, MacFORTH provides simple low-level words to access any toolbox routine.

Support

CSI maintains a technical hotline for owners of Level 1 and Level 2, and publishes a quarterly newsletter. The hotline is open during business hours and is a normal toll call.

An important part of the Mac-FORTH "safety net" is the Special Interest Group (SIG) on Compu-Serve. This is currently open to registered owners of MacFORTH only, and according to CSI about 10% of MacFORTH owners participate. As with most such SIGs, this one offers data libraries with public domain software written in MacFORTH. One data library is devoted to notices and patches for known bugs. A terminal program (with source code) is included on the Level 1 disk.

User Interface

You would expect a first-class Macintosh toolkit implementation to make very good use of the Macintosh, and, for the most part, MacFORTH does just that. Windows and menus for the MacFORTH environment are stan-

102

dard Macintosh. The text editor included in Level 1 is a subset of MacWrite, with source on disk if you want to study it or add features.

MacFORTH is so Mac oriented that rough spots in the user interface come as a bit of a shock. For example, the Forth text interpreter is plain vanilla Forth and doesn't seem to know about mousing around. Keystrokes just sort of lie there on the screen, and to edit a mistake you have to back up and retype. Gasp! Fortunately the CompuServe SIG comes to the rescue here with a handy little utility that stores the last ten commands, in editable form, in a pull-down menu. Nothing like a user's group to rescue one from primitive conditions.

Another minor annoyance is that text in the MacFORTH window is not restored after being overlaid by the editor window. When you exit the editor you are left with a blank rectangle chopped out of your dialogue with MacFORTH. This conserves memory, but is tacky. Those of you with 512K Macs will want to add screen save/restore to the editor with a few lines of code.

Developer Package, Level 3

I was only able to look at the manual for Level 3. This package is for developers who want to produce turnkey products for distribution. CSI used to charge a per-copy royalty, but the new Level 3 package replaces that policy and includes a royalty-free license agreement.

The Level 3 manual provides some excellent tutorial material and seems to be as concerned with helping you turn out a good stand-alone product as the Level 1 is with helping you make good use of the Macintosh.

The support situation is a bit different for Level 3. CSI is understandably wary of spending a lot of time on the phone with developers, so Level 3 owners are expected to use the CompuServe SIG instead of the CSI hotline when they need help. In practice, this seems to work out fine, to judge from the message traffic on the SIG (which includes heavy participation by CSI staff). If you are planning to use MacFORTH to develop a turnkey product or program for distribution,

you might buy Level 2, use the hotline and the SIG while learning MacFORTH and during program development, and then upgrade to Level 3 when you are in the final stages.

Conclusion

It is clear that the authors of MacFORTH want you to get your hands on the Macintosh and use it well. The manuals, samples, and examples all support this "Power to the People" approach. Most of the manual is about how to use the Macintosh with the tools provided by MacFORTH. The result is that, for the most part, you don't have to spend hours or days

poring over *Inside Macintosh* to get the Macintosh to do something, but can instead learn by example and hands-on experimentation. Because it is so easy to try things in Forth, using it on the Mac should lead to even greater reduction of program development time than would using Forth on a more limited machine.

DDJ

Dr. Dobb's C Tools On Disk

To complement **The Toolbook** Dr. Dobb's also offers the following programs on disk for only $19.95 each. Full source code is included and, except where indicated, both CP/M and MS or PC DOS versions are available.

Small-C Compiler

Jim Hendrix's **Small-C Compiler** is the most popular piece of software published in Dr. Dobb's 10-year history. Like a home study course in compiler design, the **Small-C Compiler** and **The Small-C Handbook** provide all you need to learn how compilers are constructed, as well as teaching the C language at its most fundamental level. **The Small-C Handbook** provides documentation for both versions; however, an addendum is recommended in addition to **The Handbook** for MS or PC DOS-specific documentation. The addendum is available for $4.95.

Small Tools: Programs for Text Processing

This package consists of programs designed to perform specific functions on text files, including: editing; formatting; sorting; merging; listing; printing; searching; changing; transliterating; copying and concatenating; encrypting and decrypting; replacing spaces with tabs and tabs with spaces; counting characters, words, or lines; and selecting printer fonts.
Documentation available for $9.95.

Small-Mac: An Assembler for Small-C

Small-Mac is a macro assembler designed to stress simplicity, portability, adaptability, and educational value. The package features simplified macro facility, C-language expression operators, descriptive error messages, object file visibility, and an externally defined machine instruction table. Included programs are: macro assembler, linkage editor, load-and-go loader, library manager, CPU configuration utility, and dump relocatable files. This program is available for CP/M systems only.
Documentation available for $9.95.

To order, send this order form with your payment to: **Dr. Dobb's Journal, 2464 Embarcadero Way, Palo Alto, CA 94303**

			QTY.		TOTAL
BOOKS					
Dr. Dobb's Toolbook`	$29.95	X		=	
Small-C Handbook	$17.95	X		=	

Check Format

	CP/M	MS or PC DOS		QTY.		TOTAL
DISKS						
Small-C Compiler			$19.95	X		=
Small Tools Text Processor			$19.95	X		=
Small Mac Assembler (for CP/M only)		▓	$19.95	X		=

			QTY.		TOTAL
MANUALS					
Small-C Compiler MSPC-DOS specific addendum to The Small-C Handbook (NOTE: The Small-C Handbook provides full documentation for the CP/M version)	$4.95	X		=	
Small Tools	$9.95	X		=	
Small Mac	$9.95	X		=	

Sub Total	$	
California residents add applicable sales tax	Tax	
Add $1.75 per item for shipping in U.S., $4.25 per item outside U.S.	Shipping	
	TOTAL $	

For CP/M system disks only, please specify one of the following formats:

☐ Kaypro
☐ Apple
☐ Zenith Z-00 DS/DD
☐ Osborne
☐ 8″ SS/SD
☐ Inquire about other formats

PAYMENT MUST ACCOMPANY YOUR ORDER

☐ Check enclosed
☐ Please charge my   ☐ VISA   ☐ M/C   ☐ Amer. Exp.

Card #_____

Exp. Date_____

Signature_____

Name_____

Address_____
(please use street address)
City_____

State_____Zip_____

Allow 6–12 Weeks for delivery   3108

Last month (September 1985, #107 page 94) John Bass described how to put together your own hard-disk interface for the Macintosh for less than $500. The following listing, np_dvr.c, is the driver routine.

Mac Toolbox Listing

```
/*
 * np_dvr.c version 1.0, July 20, 1985
 * MacSCSI non-partitioned driver routine.
 * This version was based on `he ramdisk.asm example found in
 * Aztec C 1.06 and was developed on the Mac under Aztec C.
 *
 * The changes are Copyright 1985 by John L. Bass, DMS Design
 * PO Box 1456, Cupertino, CA 95014
 * Right to use, copy, and modify this code is granted for
 * personal non-commercial use, provided that this copyright
 * disclosure remains on ALL copies. Any other use, reproduction,
 * or distribution requires the written consent of the author.
 *
 * Sources are available on diskette from Fastime, PO Box 12508
 * San Luis Obispo, Ca 93406 -- (805) 546-9141. Write for ordering
 * information on this and other Mac products.
 */

#asm
;;ts=8
;
DRVNUM    equ     5                       ;change this if conflicts with others
;
Start
          dc.w    $4f00                   ;locked, read, write, ctrl, status
          dc.w    0                       ;no delay
          dc.w    0                       ;no events
          dc.w    0                       ;no menu
          dc.w    open-Start              ;open routine
          dc.w    rdwrt-Start             ;prime routine
          dc.w    control-Start           ;control routine
          dc.w    status-Start            ;status routine
          dc.w    tst-Start               ;close routine
          dc.b    8
          dc.b    ".MacSCSI"              ;name of driver
          ds      0                       ;for alignment
;
MAXMAP              equ     640
;
drBlLen             equ     16
drNmAlBlks          equ     18
drAlBlkSiz          equ     20
drClpSiz            equ     24
drAlBlSt            equ     28
drFreeBks           equ     34
;
template
          dc.w    $d2d7                   ;signature word
          dc.l    0                       ;init time
          dc.l    0                       ;backup time
          dc.w    0                       ;attributes
          dc.w    0                       ;# of files
          dc.w    4                       ;first dir block
          dc.w    0                       ;# of dir blocks
          dc.w    0                       ;# of allocation blocks total
          dc.l    0                       ;bytes/allocation block
```

```
            dc.l    0                       ;bytes/allocation call
            dc.w    4                       ;first data block
            dc.l    1                       ;next free file #
            dc.w    0                       ;# of unused allocation blocks
            dc.b    7                       ;length of name
            dc.b    "MacSCSI"               ;name of drive
            ds      0                       ;for alignment
;
stabuf
            dc.w    0                       ;current track
            dc.b    0                       ;write protect in bit 7
            dc.b    1                       ;disk in place
            dc.b    1                       ;drive installed
            dc.b    0                       ;sidedness of drive
            dc.l    0                       ;next queue entry
            dc.w    0                       ;unused
            dc.w    DRVNUM                  ;drive number
drvref      dc.w    0                       ;driver ref num
            dc.w    0                       ;file system ID
            dc.b    0                       ;sidedness of disk
            dc.b    0                       ;needs flush flag
            dc.w    0                       ;error count
;
diskbase_
            dc.l    0                       ;base of disk memory
;
openflg
            dc.w    0                       ;once only open flag
;
open
            lea     openflg,a0              ;get once only flag
            tst.w   (a0)                    ;is it open already?
```

(Continued on next page)

```
            bne     skip                ;yes -- just exit
            st      (a0)                ;set once only flag

            move.l  #14,d0              ;size of drive queue entry
            dc.w    $a5le               ;NewPtr - system heap
            clr.w   10(a0)              ;local file system
            move.l  #DRVNUM,d0          ;drive number 5
            swap    d0
            move.w  24(al),d0           ;driver reference number
            lea     drvref,a2           ;get address of driver reference number
            move.w  d0,(a2)             ;set up status buffer
            dc.w    $a04e               ;AddDriver

            move.l  (al),-(sp)          ;push handle to resource
            dc.w    $a992               ;DetachResource

            link    a6,#-50             ;allocat the block on the stack
            move.l  sp,a0               ;setup as arg to mountvol trap
            move.w  #DRVNUM,22(a0)      ;set our drive number in block
            dc.w    $a00f               ;ask for it to be mounted
            unlk    a6                  ;clear block from stack
skip
            move.l  #0,d0               ;set non-error return
            rts                         ;exit back to caller
;
rdwrt
            movem.l d0-d7/a0-a6,-(sp)       ;save regs
```

```
        jsr      ScsiRdWr_                 ;Call the C stuff
        movem.l  (sp)+,d0-d7/a0-a6         ;restore regs
        bra      tst                       ;exit via IOdone

control
        cmp.w    #1,26(a0)                 ;is it KillIO
        bne.s    tst                       ;no, exit
        move.l   #0,d0
        rts                                ;just return
;
status
        cmp.w    #8,26(a0)                 ;is it status request??
        bne.s    tst
        movem.l  a0/al,-(sp)               ;save regs
        lea      28(a0),al                 ;get dest address
        lea      stabuf,a0
        move.l   #22,d0                    ;move 22 bytes
        dc.w     $a02e                     ;BlockMove
        movem.l  (sp)+,a0/al               ;restore regs
;
tst
        move.l   #0,d0                     ;okay return
        movem.l  d4-d7/a4-a6,-(sp)         ;save regs going into IOdone
        move.l   $8fc,a0                   ;get IOdone address
        jsr      (a0)                      ;call IOdone
        movem.l  (sp)+,d4-d7/a4-a6         ;restore them after IOdone
        rts                                ;and jump to it

        public   _Uend_,_Dorg_,_Cend_

save_
        lea      Start+(_Uend_-_Dorg_)+(_Cend_-Start),a4 ; setup baseadr
```

(Continued on next page)

```
        move.l   a0,Pbp_
        move.l   al,Dp_                           ;save arguments
        rts

restore_
        move.l   Dp_,al                          ; pass DCEptr
        rts
#endasm

#include        <memory.h>
#include        <resource.h>
#include        <desk.h>
#include        <pb.h>


/*
 * varibles used by asm routines
 * Dp and Pbp are device driver arguments handled by save/restore
 * ptr and dvrarg needed by AddDriver call
 */

DCEPtr          Dp;
ParmBlkPtr      Pbp;

/*
 * Local Stuff
 */

int     DvrState;
```

```
/*
 * ScsiRdWr -- do a driver read/write function.
 */
ScsiRdWr() {
        register struct ioParam *ip;
        long    len,part;
        short   blkno;
        char    *addr;

        save();

        ip = & Pbp->u.iop;
        part = Dp->dCtlPosition & 0xFFFFFE00;
        len = (ip->ioReqCount + 511) & 0xFFFFFE00;
        addr = ip->ioBuffer;
        while(len >= 512) {
                blkno = part>>9;
                if((Pbp->ioTrap & 0xff) == 2) {
                        ScsiRead(blkno, addr);
                 } else {
                        ScsiWrite(blkno, addr);
                }
                len -= 512;
                addr += 512;
                part += 512;
        }
        ip->ioActCount = ip->ioReqCount;

        restore();                              /* exit to I/O Done */
}
```

End Listings

112

CP/M EXCHANGE

by Bob Blum

The CP/M Exchange RCP/M system is available for your use 24 hours a day, 7 days a week. Reach it by dialing (404) 449-6588.

My queue runneth over. It seems that I always have far more projects underway than I will ever have time to complete. My workmates characterize my problem as simply not being able to manage my time efficiently. I know better, I know the real reason that this phenomenon constantly haunts me.

My time is being slowly pilfered away by the software developers who continue to write programs for CP/M. It's an insidious plot. What reasonable person would continue to support an outdated operating system such as CP/M and the archaic 8-bit microprocessors it runs on? The schemers know of my weakness for trying every new program that's advertised, and they use it to their best advantage. Some are even so bold as to send me their addicting wares of slavery without even letting me call to volunteer myself for human sacrifice.

So, as you can see, I'm not at fault for falling victim to my weaknesses.

All fun aside, more and more new 8-bit CP/M software is being introduced each day. I applaud all of those who have the foresight not to abandon a strong 8-bit market. By some standards I may be misguided, but I rest assured that I am in the best of company.

In the Queue

I'm anxiously waiting for my review copy of the latest single board computer from Micro Mint. This little barn burner, the size of a mini-floppy drive, sports the HR64180, the newest 8-bit chip from Hitachi. I'm told that this device is at least as pow-

erful as an 8-Mhz Z80 and offers a much higher level of integration, including I/O ports and other features that normally require external devices. The most exciting part is its capability to address directly 512K of memory. My plans call for a review and then for porting CP/M Plus over to it, taking full advantage of all of its new advanced features. I think in a short time that a new level of performance will be established.

Even though DRI has decided to put CP/M Plus on the mature list, and will no longer actively support it, I am continuing to port it to different computers. My most recent adventure is with the Intercontinental 8-bit S-100 master board; on-board memory management and many other performance features make the ICM board a natural for CP/M Plus.

Among the top ranking features of CP/M Plus is its capability to time and date stamp files automatically at each access. This feature is especially important to me because I have the untidy habit of not religiously tracking my disk files during a heated development session. I wait for the project to conclude before trying to make good sense out of my files and then sometimes have trouble recalling the order in which I created them, and, most important, which is the most current. Until Plu*Perfect Systems began to sell DateStamper there was no software available to add this desirable feature to CP/M 2.2. I've been using DateStamper for several weeks and have found it a most useful tool and safeguard. I will be reviewing it in a future issue of this column.

Finally, my stack of things to do wouldn't be complete without a new programming tool. I have been using DSD80, a DDT-compatible debugger with many new and useful added functions. It's a very solid package

that turns the drudgery of debugging into an art form. I will be reviewing it in the future.

Encore, Encore

After producing the fastest full featured Z80 assembler, SLR Systems could do nothing for an encore but bring out the fastest fully featured 8080 assembler.

SLRMAC shares all the operational features and configuration options of its family member Z80ASM. A total of 32 permanently configurable options combined with 20 runtime command line parameters make for the ultimate in flexibility and ease of use. I found several of the configuration options to be most interesting.

A maximum error threshold count can be set that automatically aborts the assembly when reached. This option precludes resetting your computer or in some other way manually aborting an assembly if excessive errors are present.

Another option dealing with errors sets a limit on the number of errors that can be displayed on the CRT before the display is frozen. It's no longer necessary to sit poised to strike the CTL-S key combination to stop the CRT display in an attempt to catch a glimpse of an error flying by at lightning speed. Instead, the display automatically stops upon reaching the set error count and waits for the operator's signal to continue.

One other option that I particularly like is that addresses on the listing can be generated in either logical order, high byte first, or the normal low byte-high byte fashion. The mental step of swapping the address bytes slows me down as I move around in a program listing.

Command line arguments allow almost all of the permanent configuration options to be overridden and the

specification of many more options that are significant to an individual assembly.

One particularly important option deals with the type of binary file that is output from the assembly. The H option directs the output file to be in the most common .HEX format. To execute a program written in this format, of course, requires that it first be converted to a .COM file by processing it through the LOAD utility program. If the program consists of a single phase (all routines are self-contained) then another option can be used that will dramatically speed the process of converting the source program to an executable .COM file. By including the A option on the command line, SLRMAC will assemble the program in a single pass through the source file and create a .COM output file ready for execution. This option will also work for programs that contain forward references. It seems that SLR went the extra mile and included forward reference resolution logic in the assembler so that a single pass assembly is now a workable and timesaving reality.

If you're a modular programmer, as I am, then the three options dealing with relocatable output files will be of keen interest to you. Two are variations of the pseudo-standard Microsoft naming convention, while the other permits the design maximum of seven characters to be utilized. To provide advanced features not otherwise available in other binary formats, an SLR format can also be selected. This format is not only more compact than the others but it also provides for a full 16-character external name. This increased name size has proven to be very valuable to me in extremely complex systems containing many hundreds of individual modules, not to mention how much more quickly the link editing operation runs.

The restrictions pertaining to statement labels found in most assemblers have been removed from SLRMAC in most cases. It's not at all uncommon in some assemblers to be forced to follow all labels with a colon. In others, the archaic practice of requiring that only certain types of statements have labels followed by a colon is followed. SLRMAC doesn't require that any label be followed by a colon except in the case where a duplication exists between a reserved word or a macro name. In all other instances the use of the colon is entirely optional. One final word about labels in SLRMAC: a label of any length not exceeding the maximum line length of 128 characters is permitted and the beginning 16 characters are significant.

Many problems can occur when passing relocatable expressions from program to program. Most assemblers only allow the most simple uses of this type of expression. For instance, when loading the external value into a register, any type of math involving this type of expression is prohibited, which many times requires that a great deal of extra code be written. SLRMAC allows full expression usage via a special relative data type. In this way, arithmetic operations involving multiple relocatable expressions can be resolved properly at link editing time. But a word of caution is in order; the relative data type is only accepted by SLR's own link editing program.

DDJ BACK ISSUES

The standard program counter maintenance pseudo-ops are available in SLRMAC. These include ORG, CESG, DSEG, ASEG, and COMMON, to name a few. A modified form of COMMON is available to assign up to 5 additional address spaces that can be assigned beginning addresses at link editing time. This can be very important if the program being written is targeted for a machine control environment in which dedicated portions of the address space must be defined. Again, this is an extension to the standard binary output files and as such is only available in the SLR format relocatable file.

When defining data storage areas it is not at all uncommon to need a string delimiter placed at the end of, for example, a console message. The DEFZ data definition statement will automatically generate a byte containing binary zeros at the end of the quoted string. And the DC statement will automatically turn bit 7 on in the last byte of the defined string.

A full complement of conditional assembly pseudo-ops and a full macro capability are provided. Two pseudo-ops that I have found to be especially useful are IFIDN and IFDIF. IFIDN tests for equality between two strings and sets the true condition according to the result. IFDIF performs exactly the opposite function of testing for inequality between the two strings and sets the true condition according to the result.

The method used to process the MACLIB pseudo-op is unique and worthy of comment. Intel mnemonic codes were designed to be used with the 8080, and there is no provision made for the additional instructions available on the Z80. When the program is being written for the Z80 and full use of its instruction set is desired, the extended instruction mnemonics are usually added to the assembly through a macro library. This solves the problem, but at the expense of slower run time and the necessity to maintain another macro library. SLRMAC, on the other hand, directly recognizes the extended set of Z80 mnemonics. Upon encountering the MACLIB Z80 statement the extend-

ed set of mnemonics are simply turned on, resulting in no speed degradation due to the loading of a macro library and macro generation.

SLRMAC is another notable product from SLR Systems. It not only does the job that it's advertised to do but it is also free of the many petty annoyances and inaccuracies that plague so many similar products. Its many advanced features, ease of use, and blinding speed make it a must for anybody in need of an Intel mnemonic assembler.

Your Attention Please

Before progressing further with the discussion of the Ampro Little Board computer begun last month, a few words on Z80 interrupts are in order. The Z80 and its family of related peripheral support chips have an excellent built-in interrupt system. Each support chip that has more than one active element has its own internal interrupt priority scheduler and the capability to be daisy-chained into a system. This structure does not require assistance from interrupt priority schedulers in all but the biggest systems.

The main controlling element in the system, the Z80 CPU, can handle an interrupt in one of three ways depending on the setting of the interrupt mode. This fact permits the software to determine, to a large degree, how complex the interrupt system will be.

The most basic interrupt mode, mode 0, operates identically to that of the 8080 interrupt response mode and was included to provide full compatibility with that earlier and less powerful CPU. This mode is set by executing the IM 0 instruction. With this mode, the interrupting device can place any instruction on the data buss and the CPU will execute it. Thus, the interrupting device, not the memory, supplies the next instruction to be executed.

Mode 1, when enabled by the IM 1 instruction, causes the CPU to execute a restart to memory location 38h in response to an interrupt. This action is identical to that of the nonmaskable interrupt except the call location in memory is 38h instead of

66h.

Finally, mode 2, the most powerful interrupt mode, combines three system elements to resolve the effective memory address loaded into the program counter. With a single 8-bit value supplied by the interrupting peripheral device an indirect call can be made to any memory location.

With this mode the CPU forms a 16-bit memory address. The upper eight bits are supplied from the Z80 interrupt register, and the lower eight bits are supplied by the highest priority interrupting peripheral device. This 16-bit address is used as an indirect memory address pointing to a stored table of pointer words that point to the processing routine responsible for servicing the interrupt. Actually, only seven bits are required from the interrupting device because bit 0 is always zero. This structure might seem overly complex until you consider that through software the service routine tables can be changed at will, providing the ultimate in flexibility.

Interrupts and the Little Board

The Ampro Little Board computer satisfied all the requirements I set for my RBBS system except the need for a real-time clock. My first impulse was to attach one through a serial port, but I soon remembered that both ports were busy; one with the CRT and the other with the modem. The lack of any unused I/O ports stymied any chance I had of using an external clock. The only option open to me was to look to the Little Board for a solution.

Digging into the schematics of the Little Board revealed a fairly straightforward solution to my problem. The engineers responsible for designing the Little Board used the Z80 and its related support chips exclusively to minimize interface logic requirements and arranged them to allow full use of their daisy-chained interrupt capabilities. So it seemed that all I needed was to find an unused timer, or some other source of regular interrupts, and a little software to implement a real-time clock.

Fortunately, my search was a short one. Channel 2 of the Clock Timer

Circuit (CTC), used for generating the baud rates for the serial ports, was left totally unused. All that was needed to start it working as a real-time clock was the few lines of code of Listing One (page 120).

My first task was to define a table of pointer values (section 2 in the listing) addressing the interrupt processing routines. The address of the first entry of this table would later be loaded into the interrupt vector register. Following this table are the processing routines. I chose to place the clock values in low memory locations: 08h hours, 09h minutes, 0Ah seconds, and 0Bh tenths of seconds.

Next, I had to insure that the constant flow of interrupts from the CTC would not adversely affect any other time-sensitive system component such as the floppy disk controller. The only positive way to do this was to disable the interrupts (sections 3 through 6 in the listing) prior to entering the data transfer routines. In practice, this proved to be absolutely necessary because of excessive disk errors when running with interrupts enabled.

The side effect of this, however, is to destroy the clock's accuracy. Since it's unknown exactly how long the clock will be stopped during a disk data transfer, it's impractical to add in a fudge factor upon exiting the disk routines. By comparing the computer clock against a stopwatch I found that during a highly disk intensive period the computer clock would run approximately 50% slow. Under more normal circumstances the clock wasn't accurate, but was close enough for my purpose of timing how long a user was on the RBBS.

To achieve my goal of at least 95% accuracy I would have to run the floppy disk controller with interrupts enabled, but, as I already pointed out, that produces an unbearable error rate. My only choice was to modify the hardware slightly to allow both the CTC and the floppy disk controller to run with interrupts enabled. That is a project I am currently working on. Once complete, I will publish the software and other modifications in this column.

Getting back to the real-time clock

routines, the last addition I had to make to the BIOS (sections 7 and 8 in the listing) was to initialize the CTC and the interrupt register with the proper constants at cold boot time.

If you are inclined to make these changes to your machine, be sure that you are running version 2.3 of the BIOS, also known as the hard disk BIOS. Have a CTC reference manual handy to use when setting initial values.

One final comment: My system has a hard disk that ports into the Little Board through a SCSI peer buss adapter. I have found that this configuration runs completely error free because the hard disk controller is a buffered device able to transfer data while other tasks are active.

DDJ

CP/M Exchange Listing (Text begins on page 114)

```
1) Add at line number 551 after statement FDALV  EQU  50

INTRPT   SET     YES                 ; SET INTERRUPTS TO NO AS DEFAULT
TIMER    EQU     YES                 ; USE CHANNEL 2 OF CTC AS A TIMER
TYPER    EQU     NO                  ; USE BUFFERED KEYBOARD

2) Add at line number 1028 after statement WTBLEN  EQU  $-WINCH3

         IF      TIMER

;===============================================================
;
;        MODE 2 INTERRUPT VECTORS FOR THE CTC
;
;===============================================================

         ORG     ($+17) AND 0FFF0H

TIMER$INT$TABLE:

         DW      BAUD$A
         DW      BAUD$B
         DW      TIMER$RTN
         DW      DSK$INT
         ENDIF

TIMER$RTN:

         PUSH    PSW                 ;SAVE THOSE REGISTERS THAT WE DESTROY
         PUSH    H                   ;*

         LXI     H,12                ;POINT AT HUNDREDS
         INR     M                   ;INCREMENT IT
         MOV     A,M                 ;GET HUNDREDS DIGIT
         SUI     10                  ;AT LIMIT
         JNZ     TIMER$EXIT          ;NO
         MOV     M,A                 ;ZERO HUNDREDS

         DCX     H                   ;POINT AT TENS
         INR     M                   ;INCREMENT IT
         MOV     A,M                 ;GET TENS DIGIT
         SUI     10                  ;AT LIMIT
         JNZ     TIMER$EXIT          ;NO
         MOV     M,A                 ;ZERO TENS

         DCX     H                   ;POINT AT SECONDS
         INR     M                   ;INCREMENT IT
         MOV     A,M                 ;GET SECOND DIGIT
         SUI     60                  ;AT LIMIT
         JNZ     TIMER$EXIT          ;NO
         MOV     M,A                 ;ZERO SECOND

         DCX     H                   ;POINT AT MINUTES
```
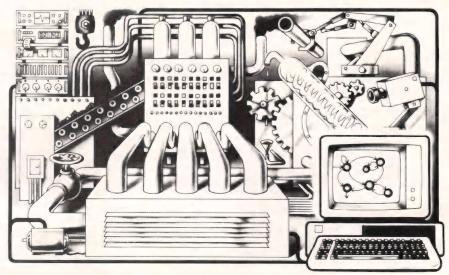
```
            INR     M                   ;INCREMENT IT
            MOV     A,M                 ;GET MINUTE DIGIT
            SUI     60                  ;AT LIMIT
            JNZ     TIMER$EXIT          ;NO
            MOV     M,A                 ;ZERO MINUTE

            DCX     H                   ;POINT AT HOURS
            INR     M                   ;INCREMENT IT
            MOV     A,M                 ;GET HOUR DIGIT
            SUI     24                  ;AT LIMIT
            JNZ     TIMER$EXIT          ;NO
            MOV     M,A                 ;ZERO HOUR

TIMER$EXIT:

            POP     H
            POP     PSW

BAUD$A:
BAUD$B:
DSK$INT:

            EI
            RETI

3) Add at line number 3144 after label WR:

            IF      INTRPT
            DI
            ENDIF
```

4) Add at line number 3159 after label WREXIT:

```
        IF      INTRPT
        EI
        ENDIF
```

5) Add at line number 3199 after label RD:

```
        IF      INTRPT
        DI
        ENDIF
```

6) Add at line number 3214 after label RDEXIT:

```
        IF      INTRPT
        EI
        ENDIF
```

7) Add at line number 4431 before label LOGMSG: and before JMP GOCPM.

```
        IF      TIMER
        XRA     A                       ;GET A ZERO
        STA     8                       ;CLEAR TIMER SAVE AREAS
        STA     9                       ;*
        STA     10                      ;*
        STA     11                      ;*
        STA     12                      ;*
        MVI     A,(LOW TIMER$INT$TABLE) AND 0F0H
        OUT     CTCA0                   ;SET INTERRUPT VECTOR LOW 8 BITS
        MVI     A,HIGH TIMER$INT$TABLE
        STAI                            ;SET INTERRUPT VECTOR HIGH 8 BITS
        IM2                             ;SET INTERRUPT MODE 2
        MVI     A,0A7H                  ;TIMER MODE - 256 PRESCALER - CONSTANT
                                        ;FOLLOWS - RESET - CONTROL
        OUT     CTCA2
        MVI     A,156                   ;TIME CONSTANT
        OUT     CTCA2
        ENDIF
```

8) Add at line number 4458 and after label LOGMD:

```
        IF      TIMER
        DB      CR,LF,'Real Time Clock Active'
        ENDIF
```

End Listing

CONTINUING THE TRADITION

DR. DOBB'S JOURNAL ANNOUNCES THE RELEASE OF BOUND VOLUME 8

Every 1983 issue together in one source

BOUND VOLUME 8

DDJ turns pro. Some of the most powerful, professional programmer's tools ever published in a magazine are in this volume. Jim Hendrix's Small C compiler. Ed Ream's RED screen editor. A microcomputer subset of the Defense Department's official programming language, Ada. C and Forth and 68000 software. Because the magazine increased in size in 1983, this volume is bigger and better that ever.

Vol. 1 1976

The material brought together in this volume chronicles the development in 1976 of Tiny BASIC as an alternative to the "finger blistering," front-panel, machine-language programming which was then the only way to do things. This is always pertinent for the bit crunching and byte saving, language design theory, home-brew computer construction and the technical history of personal computing.
Topics include: Tiny BASIC, the (very) first word on CP/M, Speech Synthesis, Floating Point Routines, Timer Routines, Building an IMSAI, and more.

Vol. 2 1977

1977 found DDJ still on the forefront. These issues offer refinements of Tiny BASIC, plus then state-of-the-art utilities, the advent of PILOT for microcomputers and a great deal of material centering around the Intel 6080, including a complete operating system. Products just becoming available for reviews were the H-8, KIM-1, MITS BASIC, Poly Basic, and NIBL.
Articles are about Lawrence Livermoore Lab's BASIC, Alpha Micro, String Handling, Cyphers, High Speed Interaction, I/O, Tiny Pilot & Turtle Graphics, many utilities, and even more. .

Vol. 3 1978

The microcomputer industry entered into its adolescence in 1978. This volume brings together the issues which began dealing with the 6502, with mass-market machines and languages to match. The authors began speaking more in terms of technique, rather than of specific implementations; because of this, they were able to continue laying the groundwork industry would

follow. These articles relate very closely to what is generally available today.
Languages covered in depth were SAM76, Pilot, Pascal, and Lisp, in addition to RAM Testers, S-100 Bus Standard Proposal, Disassemblers, Editors, and much, much more.

Vol. 4 1979

This volume heralds a wider interest in telecommunications, in algorithms, and in faster, more powerful utilities and languages, innovation is still present in every page, and more attention is paid to the best ways to use the processors which have proven longevity—primarily the 8080IZ80, 6502, and 6800. The subject matter is invaluable both as a learning tool and as a frequent source of reference.
Main subjects include: Programming Problems/ Solutions, Pascal, Information Network Proposal, Floating Point Arithmetic, 8-bit to 16-bit Conversion, Psuedo-random Sequences, and Interfacing a Micro to a Mainframe—more than ever!

Vol. 5 1980

All the ground-breaking issues from 1980 in one volume! Systems software reached a new level with the advent of CP/M, chronicled herein by Gary Kildall and others (DDJ's all-CP/M issue sold out within weeks of publication). Software portability became a subject of greater import, and DDJ published Ron Cain's immediately famous Small-C compiler—reprinted here in full.
Contents include: The Evolution of CP/M, a CP/M-Flavored C Interpreter, Ron Cain's C Compiler

for the 8080. Further with Tiny BASIC, a Syntax-Oriented Compiler Writing Language, CP/M to UCSD Pascal File Conversion, Run-time Library for the Small-C Compiler and, as always, even more!

Vol. 6 1981

1981 saw our first all-FORTH issue (now sold out), along with continuing coverage of CP/M, small-C, telecommunications, and new languages. Dave Cortesi opened "Dr. Dobb's Clinic" in 1981, beginning one of the magazine's most popular features.
Highlights: Information on PCNET, the Conference Tree, and The Electric Phone Book, writing your own compiler, a systems programming language, and Tiny BASIC for the 6809.

Vol. 7 1982

In 1982 we introduced several significant pieces of software, including the RED text editor and the Runic extensible compiler, and we continue to publish utility programs and useful algorithms. Two new columns, The CP/M Exchange and The 16-Bit Software Toolbox, were launched, and we devoted special issues to FORTH and telecommunications. Resident Intern Dave Cortesi supplied a year of "Clinic" columns while delivering his famous review of JRT Pascal and writing the first serious technical comparison of CP/M-86 and MSDOS. This was also the year we began looking forward to today's generation of microprocessors and operating systems, publishing software for the Motorola 68000 and the Zilog Z8000 as well as Unix code. And in December, we looked beyond, in the provocative essay, "Fifth-generation Computers."

Complete your reference library. Buy the entire set of Dr. Dobb's Journals from 1976 through 1983, Bound Volumes 1–8, for $195.00. That's $34.00 off the combined individual prices—a savings of almost 15%!

DDJ Classifieds

Software	Software	Utility	Hardware

TECMAR GRAPHICS LIBRARY

TEKMAR lets you do high-res graphics on your TECMAR Graphics Master. Features windowing, viewporting, clipping, axis rotation. Similar to Tekronix graphics. Includes screen dump/restore. Epson screen print, support for HP and Western Graphtec plotters. Includes three curve-fitting programs and graphics application SOURCE CODE. Requires MS-FORT 3.20, or Lahey F77L.
Price: $195.
ADVANCED SYSTEMS CONSULTANTS,
18653 Ventura Boulevard, Suite 351,
Tarzana, California 91356
(818) 407-1059

DBase II User—Converting to "C"
Try the dBx translation system— from DBase II to quality "C"; incl. translator, screen handler, & sort (w/source)— uses any file handler. For MSDOS, XENIX & UNIX Sys 5.
Desktop AI
Box 640
Norwalk, CT 06856

DBase II User—Converting to "C"?
Try the dBx translation system—from DBase II to quality "C"; incl. translator, screen handler, & sort (w/source)— uses any file handler. For MS-DOS, XENIX & UNIX sys 5.
Desktop AI
Box 640
Norwalk, CT 06856

SOFTWARE SENTINEL
A hardware key that prohibits unauthorized use of software. Its benefits: Unlimited backup copies; unbreakable; site licensing control; no floppy required with hard disk; transparent; pocketsize. Evaluation Kit available. PC compatible.
Rainbow Technologies, Inc.
17971 Skypark Circle, Suite E
Irvine, CA 92714 (714) 261-0228

Pascal's Friend
PASCAL'S FRIEND v. II contains source code for use with IBM PC Turbo Pascal: 1-2-3 style menu routines, keyboard handling, save and restore screens, read disk directory or any track and sector, system clock and calendar routines, write strings in any attribute, DOS funcion calls and MORE!
J. S. Computing
815 N. 12th St.
Suite 5
Allentown, PA 18102
(215) 821-9020

Recruiting

Personal Computer Owners
CAN EARN $1000 TO $5000
monthly selling simple services performed by their computer. Work at home in spare time. Get free list of 100 best services to offer.

Write:
C.I.L.D.G.
P. O. Box 60369
San Diego, CA 92106-8369

Dr. Dobb's Journal
is pleased to announce the
DDJ Classifieds

RATES: DISPLAY ADVERTISERS: Price per column inch $100. Ad must run in 3 consecutive issues.
LINE ADVERTISERS: Price per line $12 (35 characters per line including spaces). Minimum of 5 lines, 3 consecutive issues. Add $3 per line for boldface type. Add $25 if a background screen is desired.
DISCOUNTS: Frequency discounts for 6 consecutive ads (less 7%) and for 12 consecutive ads (less 15%).
MECHANICAL REQUIREMENTS: Camera ready art or typewritten copy only (phone orders excepted).
Display: Specify desired size, include $20 if reduction is required. Line: Specify characters in boldface, caps. Allow 6 weeks for publication.
PAYMENTS: Full payment in advance. Check, money order, Visa, M/C, AmEx.

Run this ad in _____ issues

Size _____ col x _____ inches or 1 col x _____ lines

Payments by: _____ check _____ m/o _____ Visa _____ M/C _____ AmEx

Card # _____ Exp Date _____

Signature _____

Print Name _____

Company _____

Address _____

City _____ State _____ Zip _____

Phone _____ Current Subscriber _____

Mail to or phone Shawn Horst, DDJ Classifieds, 2464 Embarcadero Way, Palo Alto, CA 94303 (415) 424-0600

The First Idea-Processor For Programmers.

The bug stops here.

FirsTime™

Has features no other editor has.

- Fast program entry through single keystroke statement generators.
- Fast editing through syntax oriented cursor movements.
- Dramatically reduced debugging time through immediate syntax checking.
- Fast development through unique programmer oriented features.
- Automatic program formatter.

FirsTime is a True Syntax Directed Editor.

As the world's most advanced syntax-directed editor, FirsTime lets you work with ideas by taking care of the low-level syntax details of your program. For example, you can generate complete statement skeletons with one keystroke. Move the cursor from one procedure to the next with one keystroke. Type in code, and it's instantly formatted (you specify the rules). Type an error, and FirsTime warns you immediately. You can continue working if you wish. Later, you can use the search-for-error command to find the error and fix it.

FirsTime Has Thorough Error Checking.

FirsTime not only checks your syntax, but also semantics. FirsTime identifies:
- Undefined variables, types and constants.
- Assignment statements with type mismatches.
- Errors in include files and macro expansions.

To Order Call: (201) 741-8188 **or write:**

SPRUCE TECHNOLOGY CORPORATION

P.O. Box 7948
Shrewsbury, NJ 07701

FirsTime Lets You Work With Ideas.

The *Zoom command* gives you a top down view of your program logic.

The *View macro command* shows the expansions of a C macro while in the editor.

The *View include file command* instantly shows you the contents of an include file.

The *Transform command* allows you to change a statement to another similar statement, for example, a *FOR* to an equivalent *WHILE*.

The *Search for next error* command allows you to find errors throughout your program.

The *Cursor movement commands* let you traverse your program by logical elements, not just characters.

FirsTime Works With Existing Files.

FirsTime works with standard ASCII files, so you can edit any existing source files.

FirsTime for Turbo Pascal	$ 74.95
FirsTime for dBase III	$125.00
FirsTime for MS-Pascal	$245.00
FirsTime for C	$295.00

In Germany, Austria and Switzerland contact:
Markt & Technik Software Verlag
Munchen, W. Germany
(089) 4613-0

FirsTime is a trademark of Spruce Technology Corporation • MS-DOS is a trademark of Microsoft Corporation • IBM is a trademark of International Business Machines Inc. • Turbo Pascal is a trademark of Borland International • dBase III is a trademark of Ashton-Tate.

127

The Best Source Debugger for C

Interactive testing: enter any C expression, statement, or function call, and it is immediately executed and the result displayed. **Direct execution** allows fast and thorough testing, makes learning C a snap.

Run-time checking: execution stops upon exception, and source code displayed. Exceptions include array reference bounds, stack overflow arithmetic or floating point error, etc. Pointers are checked for null or out of range values.

Breakpoints: *any* number of breakpoints can be set *anywhere* in your program; breakpoints are set with screen editor, not by line numbers. Breakpoints may be conditional. Single-step by statement. Interrupt execution from keyboard. Breakpoint, exception, or interruption is always shown with source code. Examine and modify data, look at stack history. Even change your program and then resume execution!

Lint-like **Compile-time checking:** argument number and sizes are checked for consistency. Never mismatch source and object code.

The best feature of all: the *fastest* C interpreter is right there when you're debugging. Make changes in seconds with the integrated screen editor. Test the changes immediately, running your program at compiled speed. Save source code for your favorite compiler, or make stand-alone executable programs. Nothing else is needed. *Instant-C* is the fastest way to get working, fully debugged C programs available today.

"We sincerely feel that *Instant-C* can have a major positive impact on programmer productivity." *Computer Language,* Feb. 85, pp. 82-83.

Instant-C is only $495. Money back for any reason in first 31 days.

Rational
Systems, Inc.

(617) 653-6194
P.O. Box 480
Natick, MA 07160

Circle **no. 7** on reader service card.

* This advertiser prefers to be contacted directly: see ad for phone number.

Advertising Sales Offices

East Coast
Walter Andrzejewski (617) 567-8361

Midwest/West Central
Michele Beaty (317) 875-0557

Northern California/Northwest
Lisa Boudreau (415) 424-0600

Southern California/AZ, NM
Beth Dudas (714) 643-9439

Advertising Director
Shawn Horst (415) 424-0600